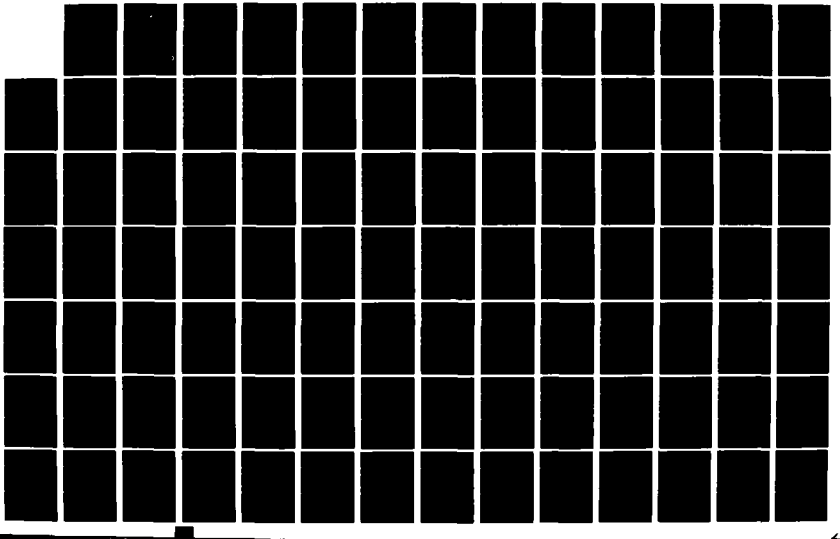
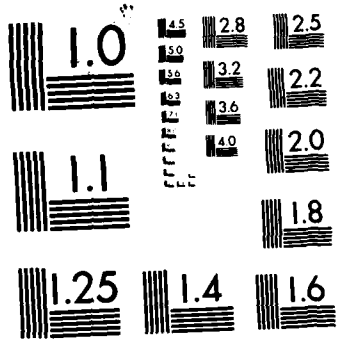
AD-A124 827 MULTIPLE MODEL ADAPTIVE ESTIMATION FOR SPACE-TIME POINT PROCESS OBSERVATIONS(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF ENGINEERING D E MEER 1/3

UNCLASSIFIED AUG 82 AFIT/DS/EE/82-2 F/G 12/1 NL





MULTIPLE MODEL ADAPTIVE ESTIMATION FOR SPACE-TIME POINT PROCESS OBSERVATIONS

DISSERTATION

AFIT/DS/EE/82-2

David E. Meer
Captain USAF

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY (ATC)
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: AFIT/DS/EE/82-2	2. GOVT ACCESSION NO.: AD-A124827
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): MULTIPLE MODEL ADAPTIVE ESTIMATION FOR SPACE-TIME POINT PROCESS OBSERVATIONS	5. TYPE OF REPORT & PERIOD COVERED: PhD Dissertation
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): David E. Meer, Capt USAF	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Air Force Institute of Technology (AFIT-EN) Wright-Patterson AFB, Ohio 45433	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: AFWL PO 82-119
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Weapons Laboratory / NTYP Kirtland AFB, NM	12. REPORT DATE: August 1982
	13. NUMBER OF PAGES: 246
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Approved for public release: IAW AFR 190-17
LYNN E. Wolaver
Dean for In...
Air Force Institute of Technology (AU)
Wright-Patterson AFB, OH 45433

18. SUPPLEMENTARY NOTES

Approved for public release; IAW AFR 190-17

Fredric C. Lynch, Major USAF, Direcor Public Affairs

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

MULTIPLE MODEL ADAPTIVE ESTIMATION	ESTIMATION
MMAE	HAZARD FUNCTION
SPACE-TIME POINT PROCESS	REGULARITY
POINT PROCESS	BAYESIAN ESTIMATION
POISSON	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This dissertation addresses the problem of estimating a vector valued stochastic process x from observations of a space-time point process which is dependent on x. The observations are corrupted by statistically independent, additive point process noise. The research is motivated by a neutral particle beam estimation and control problem in which it is desired to estimate the position of the beam from detected photo-

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

electron events. Dark current in the detector and other photon sources comprise the noise sources.

A multiple model adaptive estimator is developed in which the separate models are hypothesis sequences. The hypotheses define which observed events were due to the signal process and which were due to the noise process. The estimator provides the minimum mean squared error estimate of the underlying process. The problem is modeled on a cross product of probability spaces, and regularity conditions are defined which allow calculation of the weighting factors for the multiple model estimator. This modeling concept allows feedback from the observed events to the model, thus providing a means for control of the process. The multiple model adaptive estimator and the cross product modeling concepts are valid for a general point process signal in point process noise as long as the regularity conditions are met. The number of elemental filters in the estimator doubles as each new point process event is observed.

For the particle beam application, the elemental filters are Snyder-Fishman "firefly" filters, in which the signal process is assumed Poisson conditioned on the underlying process.

Simplifications to the full scale estimator are proposed which result in a fixed number of elemental filters. This is accomplished by considering only data within a fixed window. The data windowing is applicable to the general point process estimation problem. Simplifications which reduce the complexity of the multiple model weighting factor calculations are developed for the particle beam application. The simplifications result in a suboptimal estimator.

Monte Carlo simulations of the suboptimal estimator demonstrate that it is extremely successful at rejecting point process noise events in the measurement history, even at signal to noise count ratios as low as 0.1 and very low data rates.

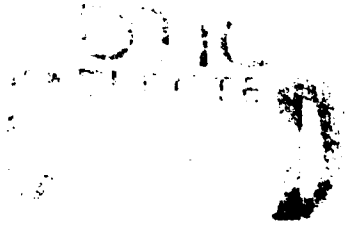
AFIT/DS/EE/82-2

MULTIPLE MODEL ADAPTIVE ESTIMATION FOR SPACE-TIME POINT PROCESS OBSERVATIONS

DISSERTATION

AFIT/DS/EE/82-2

David E. Meer
Captain USAF



AFIT/DS/EE/82-2

MULTIPLE MODEL ADAPTIVE ESTIMATION FOR SPACE-TIME POINT PROCESS OBSERVATIONS

DISSERTATION

Presented to the Faculty of the School of Engineering
of the Air Force Institute of Technology
Air University
in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

by

David E. Meer, B.S.E.E., M.S.E.E.
Captain USAF

Approved for public release; distribution unlimited.

AFIT/DS/EE/82-2

MULTIPLE MODEL ADAPTIVE ESTIMATION FOR SPACE-TIME POINT PROCESS OBSERVATIONS

by

David E. Meer, B.S.E.E., M.S.E.E.
Captain USAF

Approved:

Peter S. Maybeck 23 Sep 82

Dennis W Quinn 23 Sep 82

Albert H Moore 23 Sep 82

Ronald J Carpenter 23 Sep 82

Accepted:

J. S. Przemieniecki 24 Sep 82
Dean, School of Engineering

Preface

This research was sponsored by the Air Force Weapons Laboratory, Kirtland AFB, NM. I would like to thank the lab, and particularly Lt. Col. Raymond Zazworsky for their support and sponsorship during this project.

I wish to express my appreciation to my advisor, Dr. Peter S. Maybeck, for his technical guidance, suggestions, and encouragement during my research. His standards of excellence are matched only by his unending enthusiasm for his students and their work. A special thanks also is due my committee members, Lt. Col. Ronald Carpinella and Dr. Dennis Quinn, for their generous assistance.

Most importantly, I would like to thank my wife, Maripat, for her constant support, encouragement, and love. The words here may be mine, but the effort surely is ours. This dissertation is dedicated to you, Maripat.

Table of Contents

		Page
Preface		iii
Table of Figures		vii
Table of Notation		ix
Abstract		xii
I.	Introduction	1
	I.1 The Problem	1
	I.1.1 Problem Motivation	2
	I.1.2 Key Concepts	5
	I.2 Background Literature	5
	I.2.1 Poisson Process Estimation	6
	I.2.2 Jump Processes	6
	I.2.3 Space-Time Point Processes	7
	I.2.4 Decision Theory and Multiple Model Estimators	11
	I.3 Research Approach	13
	I.4 Summary of Remaining Chapters	14
II.	Multiple Model Adaptive Estimation	16
	II.1 Introduction	16
	II.2 The Model	17
	II.2.1 Signal	17
	II.2.2 Noise	21
	II.3 Snyder and Fishman Filter	23

II.4 Multiple Model Adaptive Estimation . . . 25

II.4.1 General MMAE 25
II.4.2 MMAE for Point Process Signal in Noise 33
II.4.3 MMAE for the Particle Beam Problem 41

II.5 Summary 59

III. Cross Product Space Model 61

III.1 Introduction 61

III.2 Doubly Stochastic Space-Time Point Process 62

III.3 Regularity 66

III.4 Implications of Regularity 74

III.4.1 Hazard Functions 75

III.5 Analytical Cross Product Space Model . . 83

III.5.1 The Observed Space, Ω 83
III.5.2 The Unobserved Space, Ω_s 85

III.6 Summary 116

IV. The Estimator 118

IV.1 Introduction 118

IV.2 The Multiple Model Weighting Factors . . 119

IV.3 Estimator Equations 135

IV.3.1 Assumptions 135
IV.3.2 The Multiple Model Adaptive Estimator 137

IV.4 Examples 144

IV.4.1 Example a: Structure Insight . 144
IV.4.2 Example b: Tracking Model . . . 146

IV.5 Summary 161

V. Filter Simplifications 162

V.1 Introduction 162

V.2 Weighting Factor Simplification 163

V.3 Limitation to a Fixed Number of Elemental Filters 171

V.3.1 Method One: Re-initialization . 172
V.3.2 Method Two: Strict Window . . . 173
V.3.3 Method Three: Best Half 175

V.4 Summary 178

VI. Simulation Results 179

VI.1 Introduction 179

VI.2 The Simulation Model 180

VI.3 General Results 182

VI.4 Parameter Sensitivity Results 189

VI.4.1 Sensitivity to g 189
VI.4.2 Sensitivity to Expected Number of Signal Events 191
VI.4.3 Sensitivity to τ 193
VI.4.4 Sensitivity to SNR 193
VI.4.5 Sensitivity to D 196
VI.4.6 Sensitivity to R 196
VI.4.7 Acquisition 199
VI.4.8 Performance Bounds 199

VI.5 Other Simulation Considerations 205

VI.6 Summary 206

VII. Conclusions and Recommendations 207

VII.1 Conclusions 207

VII.2 Recommendations 209

Bibligraphy 211

Appendix A: 216

Vita 231

Table of Figures

Figure		Page
1.	The Multiple Model Adaptive Estimator	31
2.	Hypothesis Sequences for One Measurement	35
3.	Hypothesis Sequences for Two Measurements	37
4.	Hypothesis Sequences for Three Measurements . . .	39
5.	The Spaces	67
6.	The α Coefficients	93
7.	The Detector Model	148
8.	$\bar{x}$ Storage Before and After Algorithm Step (a) .	159
9.	$\bar{x}$ Storage Before and After Algorithm Step (b) .	160
10.	Strict Window Method	174
11.	Best Half Method	177
12.	Example of True and Estimated x(t) Values for One Run	183
13.	True and Estimated Ensemble Averages for 50 Runs	185
14.	Ensemble Error Statistics for 50 Runs	186
15.	Sensitivity to Number of Monte Carlo Runs	188
16.	Estimator Performance Versus Dynamic Model Noise	190
17.	Estimator Performance Versus Expected Number of Signal Events	192
18.	Estimator Performance Versus τ	194

19. Estimator Performance Versus Signal to Noise Count Ratio 195

20. Estimator Performance Versus Depth 197

21. Estimator Performance Versus R 198

22. Unsuccessful Acquisition 200

23. Successful Acquisition 201

24. Estimator Acquisition Performance Versus R . . . 202

25. Performance Bounds 203

Table of Notation

$\bar{}$ (overbar)	Denotes vector.
$\underline{}$ (underbar)	Denotes matrix.
$\tilde{}$	Denotes random variable or process.
$\hat{}$	Denotes expected value conditioned on the measurement history.
$^{-1}$	Denotes inverse.
^{T} (superscript)	Denotes vector or matrix transpose.
$\mid$	Denotes conditioning.
A	Borel set in R^m.
A_s	Sigma field of events in Ω_s.
B	Sigma field of events in Ω.
B_t	Subsigma field of observations up to time t (the measurement history).
$\bar{c}(t)$	Control input to λ_s.
$\bar{c}'(t)$	Control input to $d\tilde{\bar{x}}(t)$ process.
$E\{\cdot\}$	Expected value.
$\underline{F}(t)$	Dynamics model matrix (nXn) for $\tilde{\bar{x}}(t)$.
$F_j(Z^{Nt})$	j^{th} order distribution function.
$\underline{G}(t)$	Dynamics model matrix (nXk) for $\tilde{\bar{x}}(t)$.
h	Dummy variable of integration for h_i.
H	The space of the hypotheses h_i.
$\underline{H}(t)$	mXn projection mapping matrix $R^n \rightarrow R^m$.
$h_j^{N_t}$	The j^{th} hypothesis sequence out of a possible 2^{Nt}.

$h_j^{N_t}(i)$	The value of hypothesis j, out of 2^{N_t} possible hypotheses, at time t_i.
$\underline{K}(t)$	Gain parameter in the Snyder-Fishman filter.
$L_t(\cdot)$	Sample function density of a point process.
N_j	The set of noise indices.
$N_t(Y)$	Number of observations in $[t_0,t), Y \subseteq R^m$
$N_{V,t}(Y)$	$\triangleq N_t(Y)-N_V(Y)$, $t_0 \leq V<t$, $Y \subseteq R^m$
N_t	Total number of observations in $[t_0,t)XR^m$.
$P(\cdot)$	Probability measure in $\mathcal{B}$.
$P_s(\cdot)$	Probability measure in A_s.
$Pr[\cdot]$	Probability of event $[\cdot]$.
R^m	Real Euclidean m space.
$\underline{R}(t)$	Dispersion matrix of $\lambda_s(\cdot)$, (mXm).
$\bar{r}_i$	The i^{th} observation location vector $(\bar{r}_i = \bar{r}(t_i))$.
$\bar{r}(t)$	Spatial measurement vector (m dimensional) at time t.
$\bar{\pi}$	Dummy variable for integration of $\bar{r}$.
S_j	The set of signal indices.
T	Borel set in $[t_0,\infty)$.
t	Time.
t_i	The i^{th} observation time.
$\underset{\sim}{\bar{u}}(t)$	Vector Wiener process (k dimensional).
t_0	Initial or starting time.

$\hat{\bar{x}}_j^{N_t}(t)$	The estimate of $\bar{x}(t)$ given the measurement history ($\mathcal{B}_t$) and the j^{th} hypothesis sequence.
$\bar{x}(t)$	n dimensional vector to be estimated.
$\hat{\bar{x}}(t)$	The estimate of $\bar{x}(t)$, given the measurement history.
Z^{Nt}	Measurement history through time t.
$\bar{z}(i)$	The measurement at t_i.
$\alpha_0(t),\alpha_1(t)$	Modeling coefficients for $\lambda_n(\cdot)$ and $\lambda_s(\cdot)$, respectively.
$\Lambda(t)$	Amplitude factor of Gaussian shaped rate parameter.
$\lambda(\cdot)$	Rate parameter for a Poisson process.
$\lambda_s(\cdot)$	Signal Poisson process rate parameter.
$\lambda_n(\cdot)$	Noise Poisson process rate parameter.
$\bar{\xi}$	Dummy variable for integration of $\bar{x}$.
$\hat{\underline{\Sigma}}(t)$	Covariance of the Snyder-Fishman filter.
$\phi(\cdot)$	Hazard function of a point process.
Ω	The observed space.
Ω_s	The unobserved space.
ω	An event in Ω .
ω_s	An event in Ω_s .

Abstract

This dissertation addresses the problem of estimating a vector-valued stochastic process $\bar{x}$ from observations of a space-time point process which is dependent on $\bar{x}$. The observations are corrupted by statistically independent, additive point process noise. The research is motivated by a neutral particle beam estimation and control problem in which it is desired to estimate the position of the beam from detected photo-electron events. Dark current in the detector and other photon sources comprise the noise sources.

A multiple model adaptive estimator is developed in which the separate models are hypothesis sequences. The hypotheses define which observed events were due to the signal process and which were due to the noise process. The estimator provides the minimum mean squared error estimate of the underlying process. The problem is modeled on a cross product of probability spaces, and regularity conditions are defined which allow calculation of the weighting factors for the multiple model estimator. This modeling concept allows feedback from the observed events to the model, thus providing a means for control of the

process. The multiple model adaptive estimator and the cross product modeling concepts are valid for a general point process signal in point process noise as long as the regularity conditions are met. The number of elemental filters in the estimator doubles as each new point process event is observed.

For the particle beam application, the elemental filters are Snyder-Fishman "firefly" filters, in which the signal process is assumed Poisson conditioned on the underlying process.

Simplifications to the full scale estimator are proposed which result in a fixed number of elemental filters. This is accomplished by considering only data within a fixed window. The data windowing is applicable to the general point process estimation problem. Simplifications which reduce the complexity of the multiple model weighting factor calculations are developed for the particle beam application. The simplifications result in a suboptimal estimator.

Monte Carlo simulations of the suboptimal estimator demonstrate that it is extremely successful at rejecting point process noise events in the measurement history, even at signal to noise count ratios as low as 0.1 and very low data rates.

MULTIPLE MODEL ADAPTIVE ESTIMATION FOR SPACE-TIME POINT PROCESS OBSERVATIONS

I. Introduction

I.1 The Problem

The problem addressed by this research is one of estimating parameters of an underlying stochastic process from observations of a point process where the point process is dependent on the underlying process and the observations are corrupted by point process noise. A second, closely related problem is that of allowing feedback control for a system in which observations of a point process signal are corrupted by point process noise. This will provide a method for investigating the optimal stochastic adaptive controller for the system.

The major contribution of this research is a method for developing an estimator for the above mentioned point process signal in point process noise environment. The method allows feedback to the model from the observations thus providing a means for control. This method is used to develop the estimator for the neutral particle beam pointing and tracking problem which motivated this research.

I.1.1 Problem Motivation. This research is motivated in part by problems in neutral particle beam pointing and tracking currently being investigated at the Ballistic Missile Defense Advanced Technology Center, Huntsville, Alabama and the Air Force Weapons Laboratory, Kirtland AFB, New Mexico. Their goal is not only to estimate the position and direction of the beam, but to use that information in an optimal way to control the pointing of the beam.

A method for sensing the location of the neutral particle beam has been proposed in which the beam is illuminated by one or more lasers. At certain angles of intersection and particle velocities, the particle electrons absorb photons from the laser beam and attain a higher energy state. The electrons spontaneously decay to their ground energy states and radiate photons in an approximately isotropic manner. By detecting this resonant scattered light energy, the position of the beam can be inferred.

The signals resulting from detection of optical fields can be modeled by conditional Poisson (CP) random processes (Refs. 25,28,33,46). The statistics are conditioned on the rate parameter of the Poisson process, which is proportional to the intensity of the received optical envelope. Real optical detectors typically have a noise mechanism which is independent of the signal and can be modeled by another conditional Poisson process. There is a non-zero

probability that electrons will be emitted from a photodetector even in the absence of incident photons. The resulting current is called dark current. In general, there will also be noise induced by background light sources in the field of view of the detector. If statistical independence is assumed between the three processes, the resulting process is also CP with a rate parameter which is the sum of the three individual rate parameters. This is shown in Chapter II.

The conditional Poisson process model is required when the level of the received signal is so low that individual photo-electron events must be counted. At higher signal rates, the observed current might be adequately modeled by a Gaussian process as is done in many communication and control type problems. Only the point process signal in point process noise case is considered in this research.

The problem then is to estimate the position of the neutral particle beam from observations of a conditional Poisson process. The observed CP process is composed of the signal (scattered resonant photons from the illuminated electrons) and noise sources. An associated problem to be considered subsequently is control of the pointing of the beam.

The observations considered here are of a space-time point process which, conditioned on the rate parameter, is Poisson. The rate parameter itself is a stochastic

process and we desire to estimate some function of the rate parameter. In this dissertation, the term space-time point process is used to denote a vector valued random process which is a mapping from the cross product of a time interval and a probability space into $[t_0,T) \times R^m$ (time cross real Euclidean m space). Each observation is of the form $(t_i,\bar{r}_i)$ where t_i is the time of occurrence for the i^{th} observation and $\bar{r}_i$ is the location in real Euclidean m space for the i^{th} observation.

The observed conditionally Poisson (CP) space-time process is composed of a CP process of interest (the signal) plus a CP noise process. The noise and signal processes are assumed to be statistically independent, resulting in the observed CP process.

A second sensing mechanism which may be exploitable for pointing and tracking in the neutral particle beam project is induced gamma radiation. When the beam strikes a target, gamma radiation is produced. This can be observed by an estimator/controller and used to direct the beam. When the signal rates are low so that point process statistics must be used to model the system adequately, the results of this research can be used for estimation and control.

Other possible applications for these techniques include tracking of missiles or satellites where the observed signal rate is low, necessitating the use of point

process models. The results presented here are not dependent on active illumination of the target so a passive tracking system with low observed signal rates could also fit the model.

I.1.2 Key Concepts. The key concepts in this research are:

1. A space-time point process signal is observed in space-time point process noise.
2. We are interested in estimating some vector which parameterizes the signal process.
3. We want to allow feedback from the observations in the model in order to provide a means of control.

For the neutral particle beam problem:

4. Both the signal and noise processes are modeled as Poisson processes, conditioned on knowledge of the respective (perhaps random) rate parameters.
5. The signal and noise processes are assumed statistically independent.

I.2 Background Literature

The literature applicable to this research can be divided into several overlapping categories. Each category addresses a portion of the beam estimation and control problem. The categories are: Poisson process estimation,

jump processes, space-time point process estimation and control, and multiple model adaptive estimation and control methods.

I.2.1 Poisson Process Estimation. References 13,25, 33,34, and 46 contain several examples of estimation for processes modeled by Poisson statistics. Snyder (Ref. 46), in particular, provides the requirements for modeling a point process with Poisson statistics and presents many useful probability densities and distributions for Poisson processes.

Most of the examples in these references are oriented towards communication type problems. In these examples, a time sequence of point events is observed and the rate parameter of the process is estimated. A second similar problem is that of estimating the presence of an on-off keyed signal in noise. These communication problems typically do not include any spatial observations of the process; however, the forms of the probability densities are analogous to those developed in this research for space-time Poisson process.

I.2.2 Jump Processes. One method of including the spatial nature of the observed process is to model it as a jump process (Refs. 8,41,42,49,50). In general, a jump process is one in which point events occur randomly in time and there is a value or weight associated with each observed

event. An example of a system which might be modeled by a jump process is urban vehicle traffic in which sensors measure time of arrival and the jump value might be speed or direction. For the problem at hand, we might model the system as a jump process in which the value (or weight) of the jump is the m dimensional spatial location of the observation. Vaca and Tretter (Ref. 49) discuss optimal estimation for the traffic example but no noise sources are considered.

Segall and Kailath (Ref. 42) consider modeling of randomly modulated jump processes. This model addresses our goal of inferring information about the observed signal point process; however, they approach noise as either an additive white Gaussian source or as an additive point source in a binary detection type of problem. Jump process models address some portions of our point process signal in point process noise problem, but the model does not fit all aspects and jump processes are not considered further in this research.

I.2.3 Space-Time Point Processes. The modeling of a system as a space-time point process (Refs. 11,12,35,38,45,46,47) is very applicable to this problem. In particular, the basic definitions and tools for statistical inference for space-time point processes are developed in Fishman (Ref. 11) and Fishman and Snyder (Ref. 12). Each observation of a space-time point process

consists of a time of event occurrence and a spatial coordinate of the event. In this research, the spatial coordinate is an element of real Euclidean m space and the i^{th} measurement consists of the pair

$$(t_i, \bar{r}_i) \varepsilon [t_0, T) \times R^m$$

where

t_i	is the time of occurrence
$\bar{r}_i$	is the spatial vector of the event
$[t_0, T)$	is the time interval of the observations
R^m	is real Euclidean m space.

If we let N_t be the number of observed point process events in the interval $[t_0, T)$, then the measurement history can be defined as

$$Z^{N_t} \triangleq \{(t_1, \bar{r}_1), \ \ldots \ , (t_{N_t}, \bar{r}_{N_t})\}$$

Snyder and Fishman (Ref. 47) present a conceptually pleasing motivation for this model and develop the associated estimator for the case when no noise is present in the observations. They pose the problem as one of tracking the centroid of a swarm of fireflies. The swarm is assumed to have a Gaussian shaped density in real Euclidean

m space and its centroid is assumed to move in real Euclidean m space as a linear function of the output of a linear n dimensional dynamical system driven by a standard Wiener process. The n dimensional output is Markov-1, but the motion of the centroid is not necessarily Markov-1. The observer is allowed to view the flashes of the fireflies and the measurements consist of the flash occurrence times and locations in m space. Given the centroid of the Gaussian shaped swarm, the flashes are assumed to occur as a Poisson process and the Gaussian shaped swarm corresponds to the rate parameter of the space-time conditionally Poisson process. They show that the estimate of the centroid is Gaussian and the structure of the estimator is analogous to a Kalman filter. There are propagation and update phases of the estimator and there is a residual term in the estimator structure similar to that of the Kalman filter. The updates occur at the event times rather than at some a priori chosen sample times. Although this model and estimator included no noise sources, it forms a basis for this research.

Snyder, Rhodes, and Hoversten (Ref. 48) extended the usefulness of this model and estimator with the demonstration of a separation theorem. They showed that, for the "firefly" tracking problem, the optimum stochastic controller is decomposable into an independently designed estimator and the linear deterministic optimal controller. The estimator is the Snyder and Fishman filter (Ref. 47) and

the linear control law is the deterministic result obtained if the output of the n dimensional dynamical system were known exactly.

This result is important since the goal of virtually any estimator used for tracking is to control some system to maintain track. The separation theorem provides the synthesis method for attaining this control. Secondly, the separation theorem provides a simple form for the optimal controller and can provide insight into a possible separation theorem for the case of point process signal observed in point process noise. If a separation theorem is not possible, this result may still provide insight for using forced certainty equivalence (Ref. 27 vol. III:17) to generate a controller.

Santiago (Ref. 38) investigated limitations of optical trackers, including the effects of noise on the "firefly" estimator. He performed simulations on the estimator both with and without unmodeled point process noise corruption of the data. As might be expected, the estimator's performance degraded significantly when the noise was present, since the noise was unmodeled. Santiago developed some ad hoc methods to reduce the influence of the noise. One method, motivated by residual monitoring techniques in Kalman filter applications, was to ignore measurements which resulted in a residual magnitude above some predetermined threshold. Simulations showed a significant improvement in the

estimator's performance with these methods of dealing with the noise, and his results suggest that a proper theoretical development of an estimator with modeled noise could be successful. Some form of disregarding or deweighting of events suspected of being caused by noise might be a useful course of action in developing an estimator for this signal-in-noise environment.

I.2.4 Decision Theory and Multiple Model Estimators. The two topics included in this literature category concern methods for disregarding or deweighting measurements which contain little or no information about the process of interest. Investigation of these topics is motivated by the results of Santiago.

Binary decision theory methods are presented by references 4,20,21,22,39, and 44. Lainiotis (Refs. 20,21,22) discusses algorithms for adaptive estimation of both the system's structure and parameters via decisions on a binary hypothesis model as each measurement is taken. Athans, Whiting, and Gruber (Ref. 4) discuss the general binary hypothesis decision theory for estimators, two methods of incorporating the weighted data, and a specific linear Gaussian model example.

In all six of these decision theory papers, Bayesian statistics, a priori knowledge of the hypotheses, and the measurement history are used to evaluate the validity of the most recent measurement. Emphasis is placed on

incorporating or rejecting the latest measurement when received and then considering it no further (except for the implicit effect each measurement has on the estimate at subsequent time). The goal is to obtain a recursive algorithm for incorporation of the most recent measurement in order to minimize memory and calculation requirements.

This differs from the concept of multiple model adaptive estimation (MMAE) and multiple model adaptive control (MMAC) as presented in references 2,3,5,9,19,26,27 vols. II and III,31,43,51, and 52. In MMAE, separate model estimators are maintained throughout the observation time interval. Magill (Ref. 26) and Athans and Chang (Ref. 2) describe the basic MMAE method. As in most of the MMAE/MMAC papers cited, the problems under study are modeled by linear systems driven by Gaussian noises and the separate models are chosen by selection of different parameter matrices for the model. This selection of models can be made either by knowledge that only a finite number of models can exist, or more commonly by discretization of the range of the parameter values. A set of estimators ("bank of filters" which are Kalman filters in the linear Gaussian model case) is designed, one matched to each distinct model. The estimates from each filter are weighted and summed to obtain the overall estimate. The weights are determined using Bayesian statistics and any a priori statistical knowledge

of the models in a manner similar to the binary decision theory estimators.

Stability and convergence of MMAE/MMAC algorithms are discussed in references 6,14,15,16,17, and 29. Baram (Ref. 6) presents consistency and convergence results for a large class of maximum a posteriori (MAP), maximum likelihood (ML), least squares (LS), and Bayesian estimators through the use of information metrics. Hawkes and Moore (Ref. 15) use similar methods to show that for a MMA estimator with a finite number of models, the weighting coefficient converges almost surely to one for the model closest to the true parameter.

I.3 Research Approach

The approach taken in this research is to develop an estimator for the space-time point process signal plus noise system using multiple model adaptive estimation techniques. For the particle beam application, Snyder and Fishman "firefly" filters are used for the individual filters in the "bank" (Ref. 47). Each assumed model (or hypothesis) is a distinct sequence specifying which observed events are noise and which observed events are signal. Once a model is specified by the hypothesis, only those measurements assumed caused by the signal process are considered by the individual filter. The overall estimate is a weighted sum of the individual filters' estimates as in the linear

Gaussian MMAE examples. The estimator is developed to admit feedback from the observations to the model. This will allow for definition of an optimal controller for the system. A covariance expression for the estimator is developed and methods for reducing the computational complexity are investigated.

I.4 Summary of Remaining Chapters

In Chapter II, the detailed signal-in-noise model for the system is defined and conditions for using Poisson statistics are specified. A brief description of MMAE is presented and the structure of the MMAE for this point process problem is developed. Weighting coefficients are developed using a probability density approach. The expressions for the coefficients are very difficult to compute, thus motivating the cross product space modeling concepts presented in Chapter III.

In Chapter III, a description of some of Fishman's statistical inference results for doubly stochastic space-time point processes is given, including a regularity definition and the implications of regularity in a point process. The beam problem is cast as a doubly stochastic space-time point process and an analytic cross product space model is developed for the problem. A regularity proof is given for this model.

In Chapter IV, the cross product space model is used to

develop the weighting factors for the multiple model adaptive estimator. The complete equations for the MMA estimator for the beam problem are presented and some example cases are presented.

The full scale estimator requires an exponentially growing amount of calculation and memory. Methods to simplify the estimator are presented in Chapter V. These methods result in suboptimal estimators.

In Chapter VI, results of Monte Carlo simulations are presented. The simulations are based on the suboptimal filter simplifications of Chapter V.

Conclusions and recommendations are presented in Chapter VII.

II. Multiple Model Adaptive Estimation

II.1 Introduction

In this chapter, the basic models for the signal and noise processes are presented and multiple model adaptive estimation for sequence hypotheses is developed. In Section II.2, the basic models for signal and noise are presented. The signal model (with no noise sources present) is the same as used by Snyder and Fishman in their filter development (Ref. 47). Their results are presented in Section II.3. In Section II.4, the concept of multiple model adaptive estimation (MMAE) is motivated in order to deal with the noise, and the general MMAE structure is developed. The MMAE concept is then applied to the general point process problem by considering each model to be described by a distinct hypothesis sequence that defines which observed events are due to noise and which are due to signal. Finally, the explicit MMAE filter for the particle beam problem is presented in which the a posteriori statistics are developed from a probability density point of view. In general, these results are difficult to compute, thus motivating Chapter III.

II.2 The Model

II.2.1 Signal. In this research, the signal source is the excited volume of neutral beam particles. The spontaneous decay of electrons results in emission of photons which may be observed by an array of photodetectors. We wish to determine the position of the beam from the observed photo-electron events.

As in Snyder and Fishman (Ref. 47), the signal is modeled as a space-time point process on $[t_0,\infty) \times R^m$. Each observation (photon detection) has associated with it a time of occurrence $t\varepsilon[t_0,\infty)$ and spatial location $\bar{r}\varepsilon R^m$. A physical detector array will result in a quantization of R^m into a finite number of possible points. This quantization and any resulting effects are not addressed in this research; the spatial measurements are allowed to assume any value in R^m (or perhaps some properly defined subspace of R^m).

Let T and A be Borel sets in $[t_0,\infty)$ and R^m respectively and let $N(T \times A)$ be the number of observed point events in $T \times A$. The number of observed events up to time t (regardless of spatial location) is defined as

$$N_t \triangleq N([t_0,t) \times R^m) \qquad (1)$$

The measurement history over the interval $[t_0,t)$ consi

of a sequence of pairs

$$(t_1,\bar{r}_1),(t_2,\bar{r}_2),\ldots,(t_{N_t},\bar{r}_{N_t}) \tag{2}$$

where t_i , $i=1,2,\ldots,N_t$ is the time of occurrence and $\bar{r}_i$ is the spatial location of the photo-electron event.

We assume that the density of the particles at time $t\varepsilon[t_0,\infty)$ and location $\bar{r}\varepsilon R^m$ is

$$\lambda_s(t,\bar{r},\underset{\sim}{\bar{x}}(t)) = \Lambda(t)\exp\left\{-\tfrac{1}{2}[\bar{r}-\underline{H}(t)\underset{\sim}{\bar{x}}(t)]^T \cdot \underline{R}^{-1}(t)[\bar{r}-\underline{H}(t)\underset{\sim}{\bar{x}}(t)]\right\} \tag{3}$$

where the under tilde denotes a random process, $\Lambda(t)$ is a known amplitude of the density function, $\underline{H}(t)$ is a known m by n projection matrix, $\underline{R}(t)$ is a known symmetric positive definite matrix for all $t\varepsilon[t_0,t_f)$, and the $\cdot(t)$ indicates dependence on time. In Chapter III, a modeling method is developed which allows these "known" parameters to be random; however, the rate parameter expression given in equation (3) will be used for the motivating particle beam problem. The vector $\underset{\sim}{\bar{x}}(t)$ is an n dimensional Gaussian output of a linear stochastic differential equation

$$d\underset{\sim}{\bar{x}}(t) = \underline{F}(t)\underset{\sim}{\bar{x}}(t)dt + \underline{G}(t)d\underset{\sim}{\bar{u}}(t)$$

$$\underset{\sim}{\bar{x}}(t_0) = \underset{\sim}{\bar{x}}_0 \qquad t \geq t_0 \tag{4}$$

where $\underline{F}(t)$ is an n by n dimensional known matrix function of time, $\underline{G}(t)$ is an n by k dimensional matrix function of time, $\underset{\sim}{\bar{u}}(t)$ is a standard k dimensional Wiener process of unit diffusion, and $\underset{\sim}{\bar{x}}_0$ is a Gaussian random variable with covariance $\underline{\Sigma}_0$ and mean $\bar{x}_0$. This definition is expanded in Section III.5 to include feedback control.

This form of $\underset{\sim}{\bar{x}}(t)$ is useful because it is descriptive of a large class of estimation and control problems, it is flexible, and it results in an estimator (described in Section II.3) which is analogous to a Kalman filter in several ways. Neither the Gaussian shaped signal rate parameter nor the linear dynamical form of $\underset{\sim}{\bar{x}}(t)$ are required for the multiple model nature of the full scale estimator. We could consider a more general form of the signal rate parameter and, perhaps, a non-linear stochastic equation to define $\underset{\sim}{\bar{x}}(t)$. With the appropriate elemental estimator for this more general model, we can still use multiple model adaptive estimation concepts for the full scale estimator.

Surfaces of constant particle density form ellipsoids in R^m and the centroid of the ellipsoids is $\underline{H}(t)\underset{\sim}{\bar{x}}(t)$.

The shape and size of the ellipsoids can change with time in a deterministic manner and the centroid moves as a Gaussian process.

We assume that:

$$\lim_{\tau,\rho\downarrow 0} \frac{\Pr[N([t,t+\tau) \times c(\bar{r},\bar{\rho}))=1 | \mathcal{B}_t, x(\sigma);\sigma\geq t_0]}{\tau\rho^m}$$

$$= \lim_{\tau,\rho\downarrow 0} \frac{\Pr[N([t,t+\tau) \times c(\bar{r},\rho)\geq 1 | \mathcal{B}_t, x(\sigma);\sigma\geq t_0]}{\tau\rho^m} \qquad (5)$$

$$= \lambda_s(t,\bar{r},\bar{x}(t))$$

where $\mathcal{B}_t$ is the sub sigma algebra of events up to time t, and where $c(\bar{r},\bar{\rho}) = [r_1,r_1+\rho_1)\times\ldots\times[r_m,r_m+\rho_m)$ is a volume in R^m. As a result, given the process $\underset{\sim}{\bar{x}}(t)$, the point events occur in $[t_0,\infty) \times R^m$ as a conditional space-time Poisson point process. The particle density $\underset{\sim}{\lambda}_s(t,\bar{r},\underset{\sim}{\bar{x}}(t))$ defined in equation (3) is the rate parameter of this conditional Poisson process.

Note that in the current description of the signal model, $\underline{R}(t)$, $\Lambda(t)$, and $\underline{H}(t)$ are time varying deterministic quantities. The model is expanded to admit randomness in each of these in Chapter III.

II.2.2 Noise. We let the noise be modeled by a space-time point process on $[t_0,\infty) \times R^m$ and assume that

$$\lim_{\tau,\rho \to 0} \frac{\Pr[N([t,t+\tau) \times c(\bar{r},\bar{\rho}))=1 | \mathcal{B}_t]}{\tau \rho^m} \tag{6}$$

$$= \lambda_n(t,\bar{r})$$

so that noise induced photo-electron events occur as a space-time Poisson process on $[t_0,\infty) \times R^m$ with rate parameter $\lambda_n(t,\bar{r})$. If we allow the rate parameter to depend on some random process $\underset{\sim}{\theta}$, then $\underset{\sim}{\lambda}_n(t,\bar{r})$ is a random process and the noise events occur as a Poisson space-time point process conditioned on knowledge of $\underset{\sim}{\theta}$.

The noise process is assumed to be statistically independent of the signal process and additive. At this point, there is no requirement to define $\underset{\sim}{\lambda}_n(t,\bar{r})$ further. We can develop the MMAE concepts with the current general description of $\underset{\sim}{\lambda}_n(t,\bar{r})$. In Chapter III, additional restrictions will be placed on $\lambda_n(t,\bar{r})$ that

$$E\{\underset{\sim}{\lambda}_n(t,\bar{r})\} < \infty$$

and

$$\int_{t_0}^{t} \int_{Y} \lambda_n(\tau,\bar{r})d\bar{r}d\tau < \infty \qquad (7)$$

where Y is R^m or a subspace of R^m to which we restrict the estimator. This constraint is necessary for the regularity proof in Chapter III and it is not physically very restrictive. The condition does imply that the observation of a noise induced event is not (probabilistically) certain. For example, the subspace Y could correspond to a finite two dimensional array of photodetectors and the noise rate parameter could be a constant.

The observed point process is composed of the sum of the signal and the noise point processes. Since the signal and noise processes are independent, the probability density function of their sum is the convolution of the individual probability density functions (Ref. 33:189), and the characteristic function of the sum is the product of the individual characteristic functions (Ref. 33:159). The characteristic function for the signal process (conditioned on $\bar{x}(t)$) is

$$\exp[\lambda_s(t,\bar{r},\bar{x}(t))(e^{j\omega}-1)]$$

and the characteristic function for the noise process (conditioned on knowledge of any uncertainties in λ_n) is

$$\exp[\lambda_n(t,\bar{r}\ )\ (e^{j\omega}-1)]$$

where, for these two expressions only, $j=\sqrt{-1}$ and ω is the frequency domain variable; this notation is used here to be consistent with the notation of Papoulis (Ref. 33). The product of these two conditional characteristic functions is the characteristic function of the sum of the processes. The form of the product is that of a conditional Poisson point process with rate parameter

$$\underset{\sim}{\lambda}(t,\bar{r},\bar{x}(t)) = \underset{\sim}{\lambda}_s(t,\bar{r},\bar{x}(t)) + \underset{\sim}{\lambda}_n(t,\bar{r}) \qquad (8)$$

II.3 Snyder and Fishman Filter

Snyder and Fishman (Ref. 47) present an estimator for the vector $\underset{\sim}{\bar{x}}(t)$ (when no noise is present) as

$$\hat{\underset{\sim}{x}}(t) = E\{\underset{\sim}{x}(t)|Z^{N_t}\} = \int_{R^m} \bar{\xi}(t) p_{\underset{\sim}{x}(t)|Z^{N_t}}(\bar{\xi}(t)|Z^{N_t})d\bar{\xi} \quad (9)$$

where $p(\bar{\xi}|Z^{N_t})$ is the conditional probability density function of $\underset{\sim}{x}(t)$ given the measurement history Z^{N_t}. The estimator is developed for the signal model defined in the previous section when there is no noise: $\lambda_n(t) = 0$. The estimator is presented in differential form as

$$d\hat{\underset{\sim}{x}}(t) = \underline{F}(t)\hat{\underset{\sim}{x}}(t) + \int_{R^m} \underline{K}(t)[\bar{r} - \underline{H}(t)\hat{\underset{\sim}{x}}(t)]N(dt \times d\bar{r}) \quad (10)$$

$$d\hat{\underline{\Sigma}}(t) = \underline{F}(t)\hat{\underline{\Sigma}}(t)dt + \hat{\underline{\Sigma}}(t)\underline{F}^T(t)dt + \underline{G}(t)\underline{G}^T(t)dt - \int_{R^m} \underline{K}(t)\underline{H}(t)\hat{\underline{\Sigma}}(t)N(dt \times d\bar{r}) \quad (11)$$

$$\underline{K}(t) = \hat{\underline{\Sigma}}(t)\underline{H}^T(t)[\underline{H}(t)\hat{\underline{\Sigma}}(t)\underline{H}^T(t) + \underline{R}(t)]^{-1} \quad (12)$$

$$\hat{\underline{\bar{x}}}(t_0) = \underline{\bar{x}}_0$$

$$\hat{\underline{\Sigma}}(t_0) = \underline{\Sigma}_0 \qquad (13)$$

where $\int \cdot N(dt \times d\bar{r})$ is a counting integral (Ref. 11). They also demonstrate that the conditional density function $p(\bar{x}(t)|Z^{N_t})$ is Gaussian.

In the expanded model, which includes an independent noise source, if we knew precisely which observed events were due to the signal process and which were due to noise, then we could, trivially, use Snyder and Fishman's filter and only consider the signal observations. We don't know which observed events are noise, but we can use a noise rejection idea through multiple model adaptive estimation techniques.

II.4 Multiple Model Adaptive Estimation

In this section, multiple model adaptive estimation is presented in general terms and then the specific MMAE equations are developed for the point process signal plus noise problem.

II.4.1 General MMAE. Let us suppose that we desire to estimate the value of some quantity $\underline{\bar{x}}(t)$ which is a random process. If we use the minimum mean square error criterion

of optimality, we can define the optimal estimate as the expected value of $\bar{x}(t)$ given the measurement history $Z^{N}t$

$$\hat{\bar{x}}(t) = E\{\underset{\sim}{\bar{x}}(t)|Z^{Nt}\} = \int \bar{\xi}(t)p(\bar{\xi}|Z^{Nt})d\bar{\xi} \qquad (14)$$

where $\bar{\xi}(t)$ is the dummy variable of integration for $\underset{\sim}{\bar{x}}(t)$ and $p(\bar{\xi}(t)|Z^{Nt})$ is the conditional probability density function of $\underset{\sim}{\bar{x}}(t)$ given the measurement history Z^{Nt}. (Note that the subscripts on the probability density function have been dropped for simplicity of notation. This convention is used in the rest of this dissertation unless the subscripts are needed for clarity.) We use the same notation to describe the measurement history here as we did for the point process model description in Section II.2 in order to maintain continuity of notation. For this general MMAE development, the measurement history Z^{Nt} is whatever measurement is appropriate to the physical problem and model under consideration. There is no implication that the general MMAE method is restricted to space-time point processes. The integration in equation (14) is over the domain of $\underset{\sim}{\bar{x}}(t)$ and we assume that the probability density function exists for the Riemann integral to have meaning.

Further suppose that we do not know how to model the process $\underset{\sim}{\bar{x}}(t)$ exactly to obtain the expression for $p(\bar{\xi}(t)|Z^{N_t})$ but that we know that the correct model is one of a finite set of possible models. (The restriction to a finite set of possible models is not necessary for MMAE in general. In this point process application, however, it will be natural to accept this type of model restriction. That course is taken in this development. Athans and Chang (Ref.2) consider linear Gaussian models which are discretized to a finite set from a possible continuum of models.)

Let there be J+1 possible models to represent the process $\underset{\sim}{\bar{x}}(t)$, where each model is represented by h_j , $j \in 0,1,2,3, \ldots, J$. Let

$$h_j \in H \tag{15}$$

where H is an appropriately defined space. In an example where the different models, or hypotheses, are represented by real matrices, the space H could be a sufficiently dimensioned real Euclidean space.

From equation (14), our definition for the optimal estimator is

$$\hat{\bar{x}}(t) = \int \bar{\xi}(t) p(\bar{\xi}(t) | Z^{N}t) d\bar{\xi} \tag{16}$$

The conditional probability density function also depends on the model h_j, so we can obtain the marginal density function from the joint density function by

$$\hat{\bar{x}}(t) = \int \bar{\xi}(t) \int_{H} p(\bar{\xi}(t), h | Z^{N}t) dh d\bar{\xi} \tag{17}$$

where h is the dummy variable for $h_j \epsilon H$. By Bayes' rule

$$\hat{\bar{x}}(t) = \int \bar{\xi}(t) \int_{H} p(\bar{\xi}(t) | h, Z^{N}t) p(h | Z^{N}t) dh d\bar{\xi} \tag{18}$$

Because we have limited the models to a finite set, the probability density function $p(h|Z^{N}t)$ is a discrete density of the form

$$p(h|Z^{Nt}) = \sum_{j=0}^{J} \Pr[h_j \text{ is correct}|Z^{Nt}]\delta(h-h_j) \qquad (19)$$

where $\delta(\cdot)$ is the Dirac delta function and $\Pr[\cdot]$ denotes probability. We can substitute equation (19) into equation (18) to get

$$\hat{\bar{x}}(t) = \int \bar{\xi}(t) \int_H p(\bar{\xi}(t)|h, Z^{Nt}) \sum_{j=0}^{J} \Pr[h_j|Z^{Nt}]\delta(h-h_j)dh d\bar{\xi} \qquad (20)$$

and by the sifting property of the delta function

$$\hat{\bar{x}}(t) = \int \bar{\xi}(t) \sum_{j=0}^{J} p(\bar{\xi}(t)|h_j, Z^{Nt})\Pr[h_j|Z^{Nt}]d\bar{\xi} \qquad (21)$$

The interchange of the integration and summation from equation (20) to equation (21) is justified by the fact that the sum is finite. By changing the order of the

integration and summation again, we obtain

$$\hat{\underset{\sim}{x}}(t) = \sum_{j=0}^{J} \Pr[h_j | Z^{Nt}] \int \underset{\sim}{\xi}(t) p(\underset{\sim}{\xi}(t) | h_j, Z^{Nt}) d\underset{\sim}{\xi} \tag{22}$$

$$= \sum_{j=0}^{J} \Pr[h_j | Z^{Nt}] E\{\underset{\sim}{x}(t) | h_j, Z^{Nt}\} \tag{23}$$

$$= \sum_{j=0}^{J} \Pr[h_j | Z^{Nt}] \hat{\underset{\sim}{x}}_j(t) \tag{24}$$

In equation (24), $\hat{\underset{\sim}{x}}_j(t)$ is the estimate of $\underset{\sim}{x}(t)$ conditioned on the measurement history and the specific model. The various densities in the above development are assumed to exist, although a parallel development can be made using probability distribution or measure theory notation. The overall structure of the estimator is shown in Figure 1.

We also desire to find the covariance for the multiple

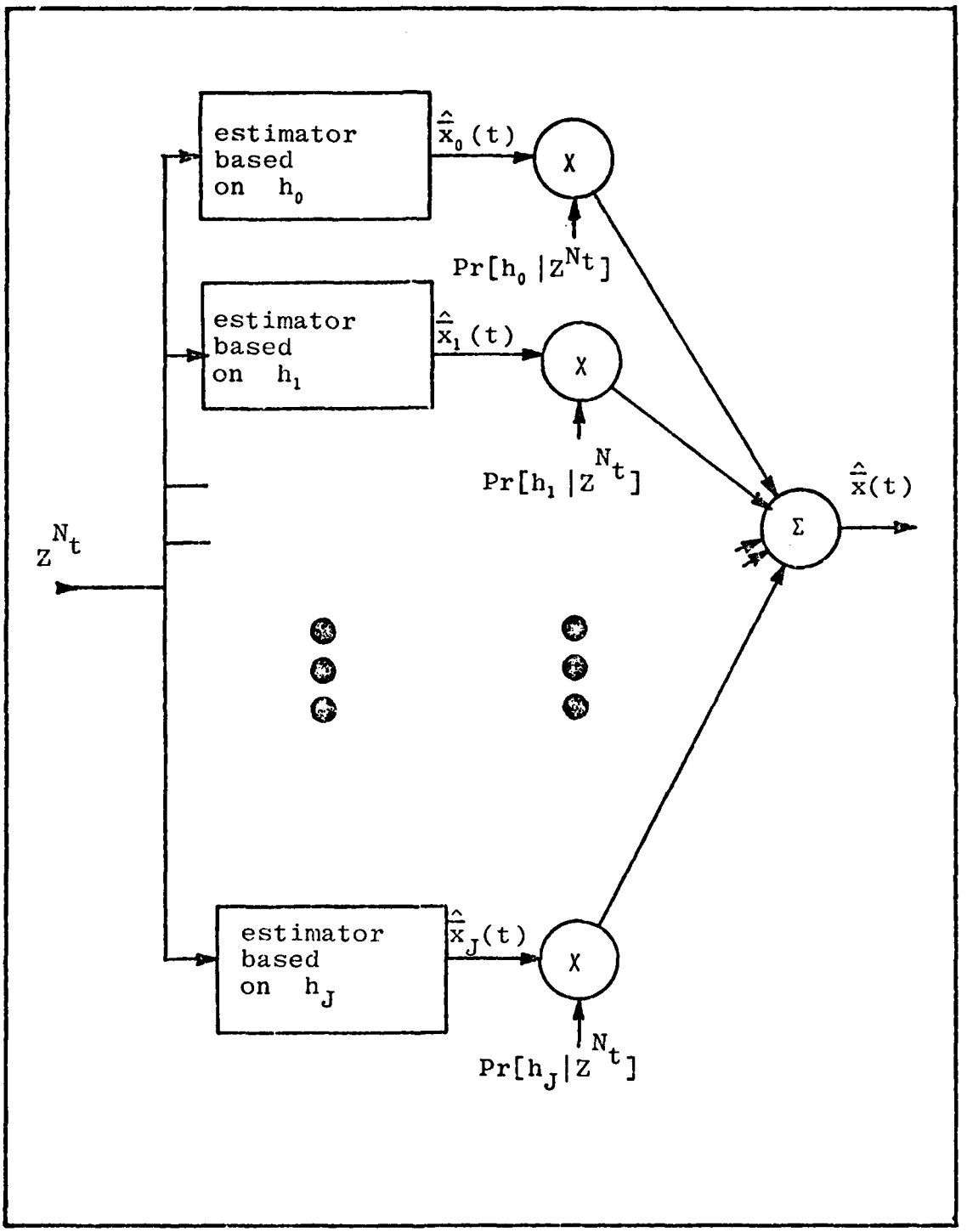


Figure 1. The Multiple Model Adaptive Estimator.

model adaptive estimator. This is useful as a measure of the estimator's performance, although it is not necessary to calculate it for the online estimator. The covariance for the full scale estimator is defined as

$$\hat{\underline{\Sigma}}(t) \triangleq E\{(\underline{\bar{x}}(t) - \hat{\bar{x}}(t))(\underline{\bar{x}}(t) - \hat{\bar{x}}(t))^T | Z^{Nt}\} \qquad (25)$$

We define the covariance of the individual estimators in the "bank" as

$$\hat{\underline{\Sigma}}_j(t) = E\left\{(\underline{\bar{x}}(t) - \hat{\bar{x}}_j(t))(\underline{\bar{x}}(t) - \hat{\bar{x}}_j(t))^T | h_j, Z^{Nt}\right\} \qquad (26)$$

The covariance of the multiple model adaptive estimator, in terms of the individual covariances and estimates, is (Ref. 2:30 and 7:420)

$$\hat{\underline{\Sigma}}(t) = \sum_{j=0}^{J} Pr[h_j | Z^{Nt}]\left[\hat{\underline{\Sigma}}_j(t) + [\hat{\bar{x}}_j(t) - \hat{\bar{x}}(t)][\hat{\bar{x}}_j(t) - \hat{\bar{x}}(t)]^T\right] \qquad (27)$$

Examples of MMAE for linear systems driven by white

Gaussian noise are given in references 2,7,26 and 27 vol. II. For these models, the individual estimators are Kalman filters, each one tuned to match the associated model hypothesis. The weighting terms, $Pr[h_j|Z^{Nt}]$, can be calculated recursively (Ref. 2:33). Except for the fact that J+1 Kalman filters must be operated simultaneously, the filter structure is computationally reasonable via distributed processing. The requirements for memory and calculation do not expand as each measurement is made.

Note that this development has assumed that only one hypothesis is correct over the observation interval. If the system is modeled such that it is allowed to switch from one hypothesis to another between measurement times, then we must calculate the a posteriori probabilities based on histories of hypotheses. This leads to an exponentially expanding number of filters in the "bank" (Ref. 7). If the switching is allowed to occur as a Markov-1 process, then the number of filters expands as the square of the number of possible states. The expanding number of hypotheses and the resulting expanding requirements for memory and calculation characterize multiple model adaptive estimation for the point process problem under consideration.

II.4.2 MMAE for Point Process Signal in Noise. With the signal and noise models described in the first section of this chapter, the observed process is conditionally

Poisson and the observations consist of a sequence of pairs, time of occurrence and spatial location, as shown in equation (2). Each observed point event ($(t,\bar{r})$ pair) is either due to the underlying signal process or to the underlying noise process. We can use this concept of a binary decision at each observed point event to construct all of the possible sequences of noise/signal events which could have produced the observed sequence.

For example, at time t_0, N_t is zero; no events have been observed. When the first point event is observed at time t_1 and location $\bar{r}_1$, we have a measurement sequence consisting of one data point for the observation interval $[t_0,t), t_1<t$. This observed event could have been caused by either the signal process or the noise process. We can represent the possible hypotheses with a tree diagram as in Figure 2.

A hypothesis sequence is denoted as $h_j^{N_t}$ where the subscript $j\varepsilon\{0,1, \ldots ,(2^{N_t})-1\}$ denotes which particular sequence is identified and the superscript N_t is the number of data points observed up to the time when the sequence $h_j^{N_t}$ is defined. When an argument is present as in $h_j^{N_t}(i)$, $i=1,2,3, \ldots N_t$, we refer to the value of the sequence at time t_i . A hypothesis sequence can be written as

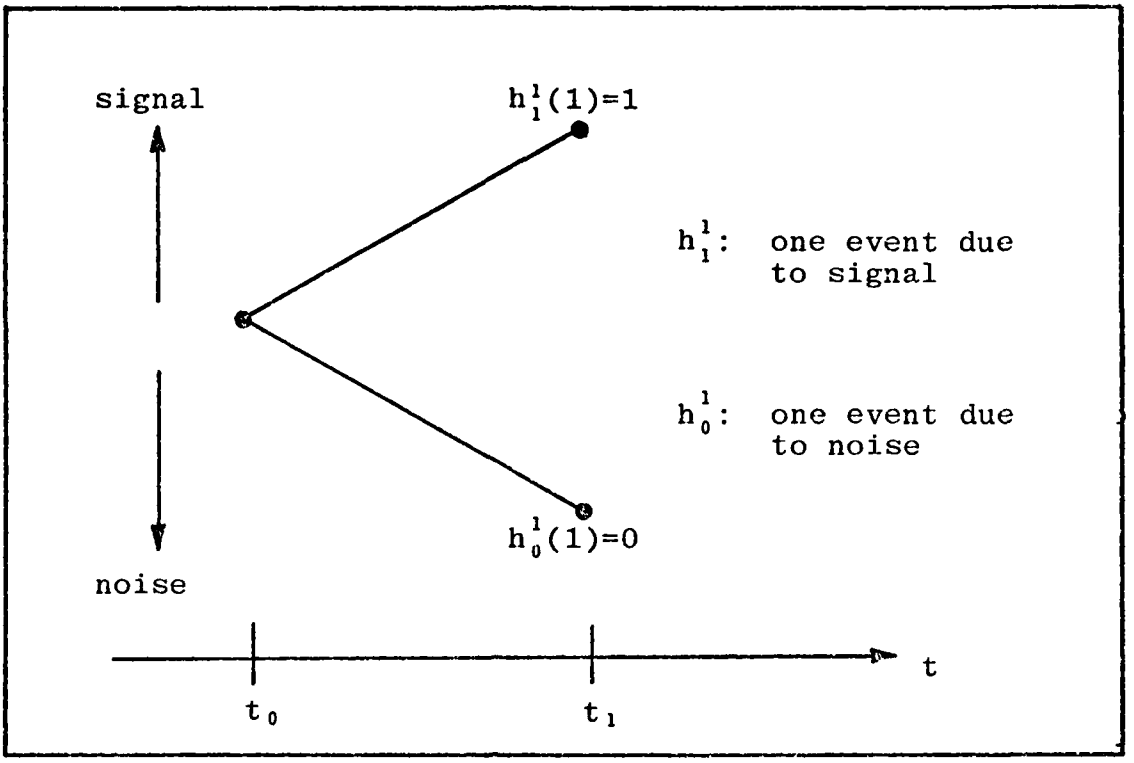


Figure 2. Hypothesis Sequences for One Measurement

$$h_j^{N_t} = \{h_j^{N_t}(1), h_j^{N_t}(2), \ldots h_j^{N_t}(N_t)\} \tag{28}$$

The notation can be understood more easily in the following examples.

In Figure 2, the hypothesis sequence h_1^1 is that the event observed at t_1 was due to signal and hypothesis sequence h_0^1 denotes that the event at t_1 was due to noise. The sequence h_1^1 is composed of the single entry

$$h_1^1 = \{h_1^1(1)\} = \{1\} \tag{29}$$

where $h_1^1(1)$ is the value of the sequence h_1^1 at time t_1. Values are assigned as

$$h_j^{Nt}(i) = \begin{cases} 1 : \text{event due to signal} \\ 0 : \text{event due to noise} \end{cases} \tag{30}$$

By the same method, the sequence h_0^1 is defined as

$$h_0^1 = \{h_0^1(1)\} = \{0\} \tag{31}$$

The upward branches on the tree denote data points which we associate with the signal process and the downward branches are for noise caused data points.

When a second data point is observed at t_2, there are

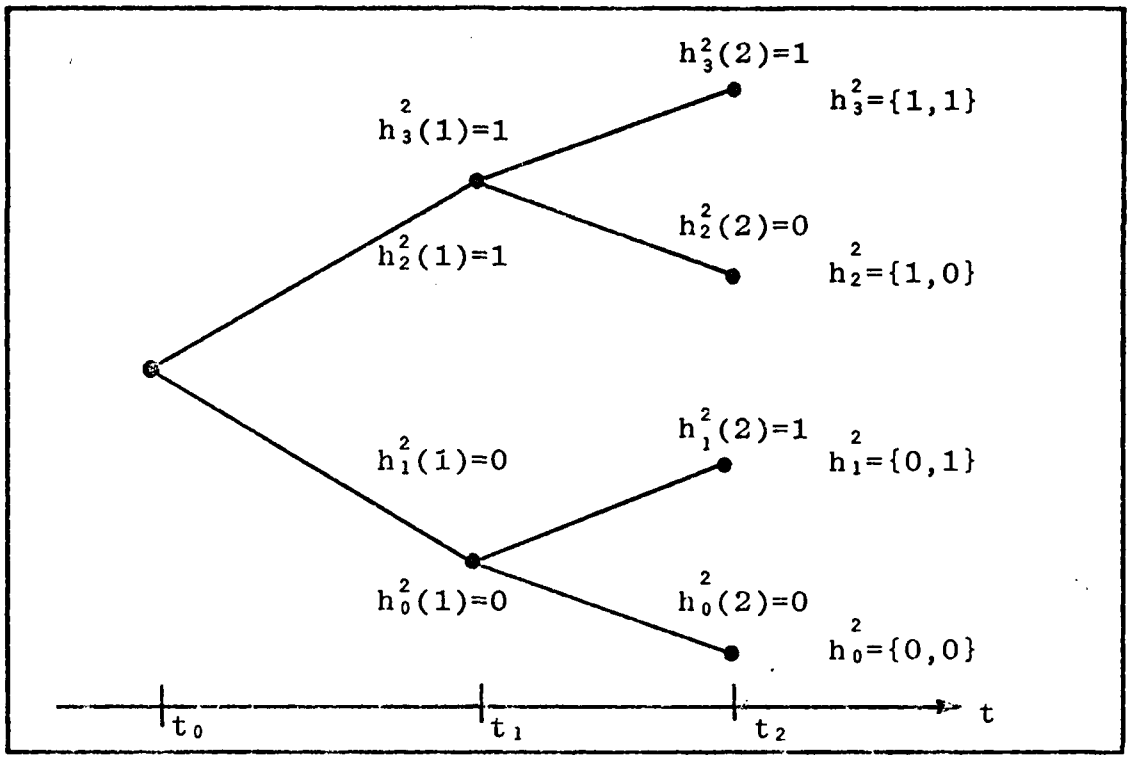


Figure 3. Hypothesis Sequences for Two Measurements

four possible sequences to describe the origination of the two data points over the interval $[t_0,t)$, $t_2<t$. These can be shown as in Figure 3.

From Figure 3, it can be seen that under hypothesis h_2^2 , the first observed event (at time t_1) is assumed to have been caused by the signal process ($h_2^2(1)=1$) . The second observed event is assumed to have been caused by the noise process ($h_2^2(2)=0$). The entire sequence can be explicitly written as

$$h_2^2 = \{h_2^2(1), h_2^2(2)\} = \{1,0\} \tag{32}$$

In a similar manner, the other three possible hypotheses are defined, for $N_t=2$, as

$$h_0^2 = \{0,0\}$$

$$h_1^2 = \{0,1\} \tag{33}$$

$$h_3^2 = \{1,1\}$$

Figure 4 shows the tree diagram for the time interval $[t_0,t)$, $t_3<t$. For clarity, only the values of one sequence are labeled.

From this example, it can be seen that for any time interval $[t_0,t)$, there are exactly 2^{Nt} possible sequences. These sequences describe all of the possible ways in which a signal process and a noise process could have caused the observed measurement history. Thus, we need only consider a finite (although growing) number of possible models as described in the last section. The close analogy to the time varying parameter case of reference 7 can now be

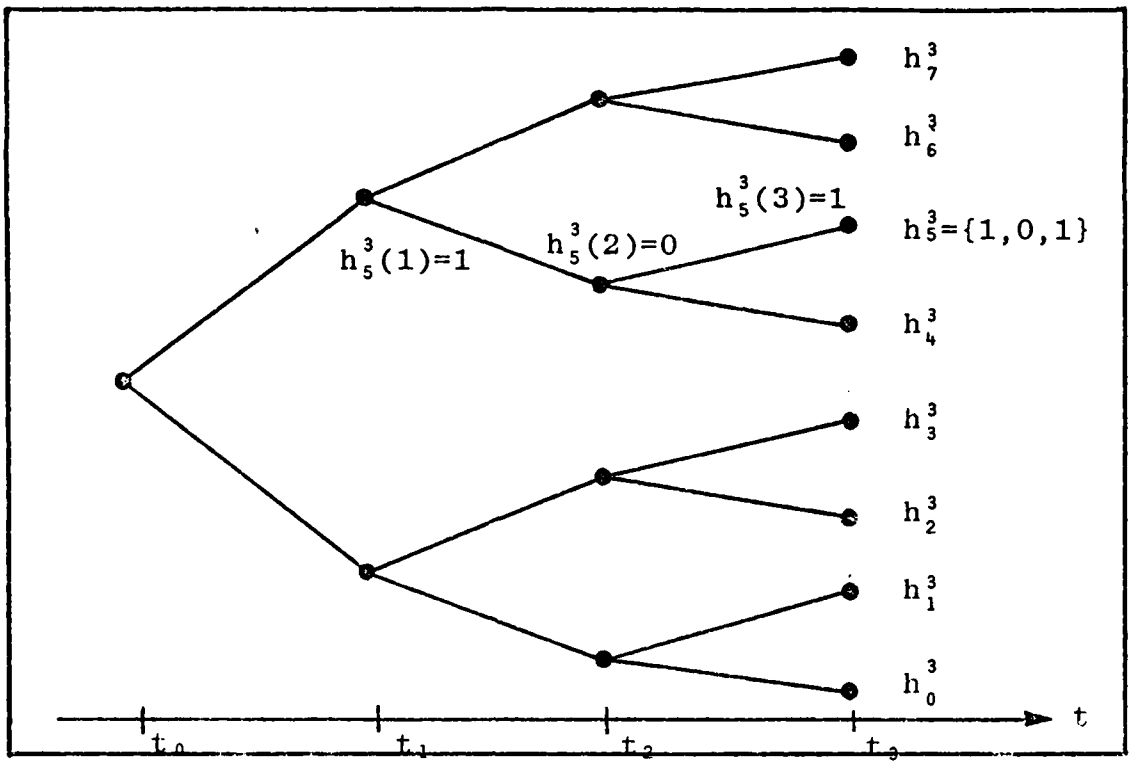


Figure 4. Hypothesis Sequences for Three Measurements

seen.

In order to use MMAE for the point process signal plus noise problem, we associate each hypothesis $h_j^{N_t}$ with a distinct model of the observed process and expand our hypothesis definition of equation (15) to

$$h_j^{N_t} \in H, \; j \in 0,1,2, \ldots , 2^{N_t}-1 \qquad (34)$$

For the time interval $[t_0,t)$, the measurement history is

$$Z^{N_t} = \{(t_1,\bar{r}_1),(t_2,\bar{r}_2), \ldots ,(t_{N_t},\bar{r}_{N_t})\} \tag{35}$$

The number of events, N_t , is implicitly included in Z^{Nt}. We can apply equation (24) to obtain the expression for the optimal estimate of $\bar{\underset{\sim}{x}}(t)$

$$\hat{\bar{x}}(t) = \sum_{j=0}^{(2^{Nt})-1} Pr[h_j{}^{N_t}|Z^{N_t}]\hat{\bar{x}}_j(t) \tag{36}$$

where

$$\hat{\bar{x}}_j(t) \triangleq E\{\bar{\underset{\sim}{x}}(t)|h_j{}^{N_t},Z^{N_t}\} \tag{37}$$

Because of the assumption that the signal and noise are independent, we can ignore all observed data points which, based on the hypothesis sequence $h_j{}^{Nt}$, are caused by noise. Equation (37) is the estimate of $\bar{\underset{\sim}{x}}(t)$ obtained by only considering those data observations for which $h_j{}^{Nt}(i)=1$ for $i \in 1,2,3,\ldots N_t$. These individual filters

are thus "tuned" to the respective hypothesis sequences.

The structure of the overall estimator is the same as shown in Figure 1, however, now the form of the individual estimators depends on the particular point process signal under consideration. For the conditionally Poisson point process signal model described in Section II.2, the individual estimators are the Snyder-Fishman filters described in Section II.3.

A significant difference between sequence hypothesis MMAE and the constant parameter linear Gaussian MMAE case described previously is that, in the sequence case, the number of filters doubles with each observed data point. There is a close analogy between sequence hypothesis MMAE for a point process model and the time varying parameter multiple model adaptive estimator. The growing number of filters places a serious computational burden on the estimator. Methods of alleviating this burden are considered in Chapter V.

II.4.3 MMAE for the Particle Beam Problem. Equation (36) defines the overall estimate of $\bar{\underset{\sim}{x}}(t)$ and Snyder and Fishman's results (equations 10-13) provide the means for evaluating the individual model estimates $\hat{\underset{\sim}{x}}_j(t)$. The only remaining term to specify analytically is $Pr[h_j^{N_t} | Z^{N_t}]$, the weighting factor.

To simplify notation, denote a single space-time observation as

$$\bar{z}(i) = (t_i, \bar{r}_i) = \begin{bmatrix} t_i \\ - - - \\ \bar{r}_i \end{bmatrix} \tag{38}$$

and the measurement history over $[t_0, t]$ as

$$Z^{N_t} = \{\bar{z}(1), \bar{z}(2), \ldots, \bar{z}(N_t)\} \tag{39}$$

Note that the sub-sigma field B_t includes all the information in Z^{N_t}.

In terms of this notation, the goal is to evaluate

$$\Pr[h_j^{N_t} | Z^{N_t}]$$

$$= \Pr[h_j^{N_t}(1), \ldots, h_j^{N_t}(N_t) | \bar{z}(1), \ldots, \bar{z}(N_t)] \tag{40}$$

The probabilities of equation (40) can be combined into a discrete probability density function. We can use Bayes' rule to obtain (Ref. 33:176)

$$\sum_{j=0}^{2^{N_t}-1} \delta(h-h_j^{N_t}) Pr[h_j^{N_t} | Z^{N_t}] = \frac{p(h, Z^{N_t})}{p(Z^{Nt})} \qquad (41)$$

The denominator of equation (41) is a probability density function evaluated at Z^{Nt}. It is the probability density of obtaining the specific realization Z^{Nt} over the interval $[t_0,t)$, $t_{N_t}<t<t_{N_t+1}$, from the space-time conditionally Poisson process with rate parameter $\lambda(t,\overline{r},\underset{\sim}{\overline{x}}(t))$ defined by equation (8). This density is termed the sample function density (sfd). Snyder (Ref. 46) develops the sample function density for a temporal (no spatial dependence) Poisson process. The brief development of the sample functior density presented here for a space-time Poisscn process follows Snyder's method.

Because the process in question is Poisson conditioned on knowledge of $\underset{\sim}{\overline{x}}(t)$ and any uncertainties in $\underset{\sim}{\lambda}_n(\cdot)$, we first consider the (trivial) case of $\underset{\sim}{\overline{x}}(t)$ and $\underset{\sim}{\lambda}_n(\cdot)$ known exactly. The random case will be discussed at the end of this section. Two additional pieces of notation will be useful for developing the sample function density. Let

$$N_t(Y) \stackrel{\Delta}{=} N([t_0,t) \times Y) \tag{42}$$

where $Y \subseteq R^m$. Thus $N_t(Y)$ is the number of space-time point events observed on $[t_0,t) \times Y$. Note, that if $Y = R^m$, then $N_t(Y)=N_t$ as defined in equation (1). For $t_0 \leq v<t$, let

$$N_{v,t}(Y) \stackrel{\Delta}{=} N_t(Y) - N_v(Y) \tag{43}$$

Since we have assumed a space-time Poisson model, $N_{v,t}(Y)$ is Poisson distributed with rate parameter $\lambda(t,\bar{r},\bar{x}(t))$.

We begin by writing the probability that the realization ζ^{N_t} of the process occurs within some small space-time volume which includes Z^{N_t}

$$\Pr[\zeta^{N_t} \varepsilon [Z^{N_t}, Z^{N_t} + \Delta Z)]$$

$$= \Pr[\tau_1 \varepsilon [t_1, t_1+\Delta t_1), \ldots, \tau_{N_t} \varepsilon [t_{N_t}, t_{N_t}+\Delta t_{N_t}),$$

$$\bar{\pi}_1 \varepsilon c(\bar{r}_1,\bar{\rho}_1), \ldots, \bar{\pi}_{N_t} \varepsilon c(\bar{r}_{N_t},\bar{\rho}_{N_t})] \tag{44}$$

where ζ^{N_t}, τ_i and $\bar{\pi}_i$ are the dummy variables for Z^{N_t}, t_i and $\bar{r}_i$ respectively, $i\varepsilon\{1,2,3, \ldots N_t\}$, and the observation interval is $[t_0,t)X R^m$. The cubes in R^m are defined as before except they are now indexed in time by i:

$$c_i(\bar{r}_i,\bar{\rho}_i) \triangleq [r_{i1},r_{i1}+\rho_{i1})X\ldots X[r_{im},r_{im}+\rho_{im}) \triangleq c_i \qquad (45)$$

Equation (44) can be equivalently written as

$$\Pr[\zeta^{N_t}\varepsilon[Z^{N_t},Z^{N_t}+\Delta Z)] =$$

$$\Pr[N_{t_0,t_1}(R^m) = 0,\ N_{t_1,t_1+\Delta t_1}(\bar{r}_1\varepsilon c_1) = 1, N_{t_1,t_1+\Delta t_1}(\bar{r}_1\not\varepsilon c_1)=0,$$

$$N_{t_1+\Delta t_1,t_2}(R^m) = 0,\ N_{t_2,t_2+\Delta t_2}(\bar{r}_2\varepsilon c_2) = 1, N_{t_2,t_2+\Delta t_2}(r_2\not\varepsilon c_2)=0,$$

$$\vdots \qquad (46)$$

$$, N_{t_{N_t},t_{N_t}+\Delta t_{N_t}}(\bar{r}_{N_t}\not\varepsilon c_{N_t})=0, N_{t_{N_t}+\Delta t_{N_t},t}(R^m)=0]$$

Because $N_{v,t}(Y)$ is distributed as a Poisson process, we can use the independent increment property to factor the right hand side of equation (46) as

$$\Pr[N_{t_0,t_1}(R^m)=0]\Pr[N_{t_1,t_1+\Delta t_1}(\bar{r}_1 \varepsilon c_1)=1]\Pr[N_{t_1,t_1+\Delta t_1}(\bar{r}_1 \notin c_1)=0]$$

$$\cdot\Pr[N_{t_1+\Delta t_1,t_2}(R^m)=0]\ldots\Pr[N_{t_{N_t},t_{N_t}+\Delta t_{N_t}}(\bar{r}_{N_t} \notin c_{N_t})=0] \quad (47)$$

$$\cdot\Pr[N_{t_{N_t}+\Delta t_{N_t},t}(R^m) = 0]$$

Each of these probability terms can be written in integral form and the terms collected to arrive at

$$\Pr[\zeta^{N_t} \varepsilon [Z^{N_t}, Z^{N_t}+\Delta Z)]$$

$$= \prod_{i=1}^{N_t} \left[\int_{t_i}^{t_i+\Delta t_i} \int_{\bar{r}_i \varepsilon c_i} \lambda(\tau,\bar{\pi},\bar{x}(\tau)) d\bar{\pi} d\tau \right] \quad (48)$$

$$\cdot \exp \left[-\int_{t_0}^{t} \int_{R^m} \lambda(\tau,\bar{\pi},\bar{x}(\tau)) d\bar{\pi} d\tau \right]$$

With this expression, we can now develop the probability density function. The sample function density is defined as the limit of the probability (defined by equation (48)) divided by the incremental space-time volume as the space-time volume goes to zero (Ref. 46:58):

$$p(Z^{N_t}) = \lim_{\substack{\Delta t_i \to 0 \\ \rho_{ij} \to 0}} \frac{\prod_{i=1}^{N_t} \int_{t_i}^{t_i+\Delta t_i} \int_{\bar{r}_i \varepsilon c_i} \lambda d\bar{\pi} d\tau \exp\left[-\int_{t_0}^{t} \int_{R^m} \lambda d\bar{\pi} d\tau \right]}{\prod_{i=1}^{N_t} \left[\Delta t_i \prod_{j=1}^{m} \rho_{ij} \right]} \qquad (49)$$

where the arguments of $\lambda(t,\bar{r},\bar{x}(t))$ have been dropped for simplicity, and ρ_{ij} is the j^{th} element of $\bar{\rho}$ as in (45).

As written, the limit in equation (49) is indeterminate; however, by repeated use of l'Hospital's rule and Leibnitz's rule, the limit can be evaluated as

$$p(Z^{N_t}) = \prod_{i=1}^{N_t} \lambda(t_i, \bar{r}_i, \bar{x}(t_i))$$

$$\cdot \exp\left[-\int_{t_0}^{t}\int_{R^m} \lambda(\tau, \bar{\hbar}, \bar{x}(\tau))\, d\bar{\hbar}\, d\tau\right] \qquad (50)$$

$$\text{for } t_{N_t} \leq t < t_{N_t+1}$$

Note that $\bar{x}(t)$ is assumed known in equation (50). For subsequent use, we note that equation (50) can be factored into the recursion

$$p(Z^{N_t}) = \lambda(t_{N_t}, \bar{r}_{N_t}, \bar{x}(t_{N_t})) \cdot \exp\left[-\int_{t_{N_t-1}}^{t}\int_{R^m} \lambda\, d\bar{\hbar}\, d\tau\right] \cdot p(Z^{N_t-1}) \qquad (51)$$

$$\text{for } t_{N_t} \leq t < t_{N_t+1}$$

where $p(Z^{N_t-1})$ is evaluated at time t_{N_t-1}.

We now turn to evaluation of the numerator in equation (41) for values of $h=h_j^{N_t}$. By the use of Bayes' rule, we can write

$$p(h_j^{N_t}, Z^{N_t}) = p(h_j^{N_t}(N_t), \ldots, h_j^{N_t}(1), \bar{z}(N_t), \ldots, \bar{z}(1)) \quad (52)$$

$$= p(h_j^{N_t}(N_t), \bar{z}(N_t) | h_j^{N_t}(N_t-1), \ldots, h_j^{N_t}(1), \bar{z}(N_t-1), \ldots, \bar{z}(1)) \cdot$$

$$p(h_j^{N_t}(N_t-1), \ldots, h_j^{N_t}(1), \bar{z}(N_t-1), \ldots, \bar{z}(1)) \quad (53)$$

$$= p(h_j^{N_t}(N_t), \bar{z}(N_t) | h_{j'}^{N_t-1}, Z^{N_t-1}) p(h_{j'}^{N_t-1}, Z^{N_t-1}) \quad (54)$$

At this point, the j' notation requires some explanation. The sequence $h_j^{N_t}$ is defined as

$$h_j^{N_t} = \begin{cases} \left\{ h_{j'}^{N_t-1}, 0 \right\} & : \text{if } h_j^{N_t}(N_t) = 0 \\ \left\{ h_{j'}^{N_t-1}, 1 \right\} & : \text{if } h_j^{N_t}(N_t) = 1 \end{cases} \quad (55)$$

The variable j is an index for keeping track of the sequences. As each new data point is observed, the number of possible sequences doubles and the numbering (indexing) of nearly every sequence changes as shown in Figures 2,3 and 4. The only sequence which does not change index value is the "all noise" sequence. Therefore, the value of j is almost never equal to the value of j' in equation (54).

This complicated notation is necessary due to the expanding number of sequences. The concept can be stated clearly in words as:

> $h_{j'}^{N_t-1}$ is the sequence (out of 2^{N_t-1} possible sequences) which is concatenated with the sequence value $h_j^{N_t}(N_t)$ to obtain the sequence $h_j^{N_t}$.

The actual values of the indices j and j' are not important to the MMAE problem; it is only necessary that the correct sequence be identified. Obviously, the value of the indices is important to the implementation of the estimator.

We can now combine equations (54) and (51) to show that

$$\Pr[h_j^{N_t} | Z^{N_t}] = \frac{p(h_j^{N_t}, Z^{N_t})}{p(Z^{N_t})}$$

$$= \frac{p(h_j^{N_t}(N_t), \bar{z}(N_t) | h_{j^-}^{N_t-1}, Z^{N_t-1})}{\lambda(t_{N_t}, \bar{r}_{N_t}, \bar{x}(t_{N_t}))\exp\left[-\int_{t_{N_t-1}}^{t_{N_t}} \int_{R^m} \lambda d\bar{r} d\tau\right]} \cdot \frac{p(h_{j^-}^{N_t-1}, Z^{N_t-1})}{p(Z^{N_t-1})} \quad (56)$$

where the upper limit of integration on the time integral has been changed from t to t_{N_t} to reflect the fact that we are interested in evaluating $\Pr[h_j^{N_t} | Z^{N_t}]$ immediately after having observed the point event at t_{N_t}.

From Bayes' rule in the form of equation (41) it can be seen that equation (56) is recursive :

$$\Pr[h_j^{N_t} | Z^{N_t}] =$$

$$\frac{p(h_j^{N_t}(N_t), \bar{z}(N_t) | h_{j^-}^{N_t-1}, Z^{N_t-1})}{\lambda(t_{N_t}, \bar{r}_{N_t}, \bar{x}(t_{N_t}))\exp\left[-\int_{t_{N_t-1}}^{t_{N_t}} \int_{R^m} \lambda d\bar{r} d\tau\right]} \quad (57)$$

$$\cdot \Pr[h_{j^-}^{N_t-1} | Z^{N_t-1}]$$

The one remaining term to specify for completion of

this MMAE example is the numerator

$$p(h_j^{N_t}(N_t), \bar{z}(N_t) | h_{j'}^{N_t-1}, Z^{N_t-1}) \tag{58}$$

The method for evaluating this mixed (discrete and continuous) density is similar to that used to develop the sample function density. There are two possible values for $h_j^{N_t}(N_t)$, 0 or 1. Consider $h_j^{N_t}(N_t) = 0$ first:

$$Pr\left[h_j^{N_t}(N_t) = 0, \bar{z}(N_t) \varepsilon [t_{N_t}, t_{N_t}+\Delta t) \times c_{N_t} | h_{j'}^{N_t-1}, Z^{N_t-1}\right]$$

$$= Pr\Big[\text{no signal event in } [t_{N_t-1}, t_{N_t}) \text{ and only one noise event in } [t_{N_t-1}, t_{N_t}+\Delta t) \text{ at } [t_{N_t}, t_{N_t}+\Delta t) \times c_{N_t} | h_{j'}^{N_t-1}, Z^{N_t-1}\Big] \tag{59}$$

In equation (59), the conditioning is on a specific value of $h_{j'}^{N_t-1}$ and an observed realization of the

measurement history Z^{N_t-1} . Since these are both values (or realizations as opposed to functional forms dependent on λ_s or λ_n), and since the signal and noise processes are independent, we can factor equation (59) as

$$\Pr[\cdot] = \Pr\left[\text{no signal event in } [t_{N_t-1}, t_{N_t}) \,|\, h_j^{N_t-1}, Z^{N_t-1}\right]$$

$$\cdot \Pr\Big[\text{only one noise event in } [t_{N_t-1}, t_{N_t}+\Delta t) \text{ at } [t_{N_t}, t_{N_t}+\Delta t) \times c_{N_t} \,|\, h_j^{N_t-1}, Z^{N_t-1}\Big] \tag{60}$$

(If we were not given values, the joint conditional probability density is not, in general, factorable.) By the independent increment property of the signal and noise Poisson processes, the conditioning can be dropped resulting in

$$\Pr[\cdot] = \Pr\left[\text{no signal event in } [t_{N_t-1}, t_{N_t})\right]$$

$$\cdot \Pr\Big[\text{only one noise event in } [t_{N_t-1}, t_{N_t}+\Delta t) \text{ at } [t_{N_t}, t_{N_t}+\Delta t) \times c_{N_t}\Big] \tag{61}$$

$$= \exp\left[- \int_{t_{N_t-1}}^{t_{N_t}} \int_{R^m} \lambda_s d\bar{r} d\tau \right] \int_{t_{N_t}}^{t_{N_t}+\Delta t} \int_{\bar{r}_{N_t} \epsilon c_{N_t}} \lambda_n d\bar{r} d\tau$$

$$\cdot \exp\left[- \int_{t_{N_t-1}}^{t_{N_t}} \int_{R^m} \lambda_n d\bar{r} d\tau \right] \qquad (62)$$

where the arguments of λ_s and λ_n have been dropped for simplicity. We can obtain the density for equation (58) when $h_j^{N_t}(N_t)=0$ (as we did for the sample function density development) by a limiting method :

$$p\left(h_j^{N_t}(N_t)=0, \bar{z}(N_t) \mid h_j^{N_t-1}, Z^{N_t-1}\right) =$$

$$\lim_{\substack{\Delta t \to 0 \\ \bar{\rho}_{N_t} \to 0}} \frac{\exp\left[- \int_{t_{N_t-1}}^{t_{N_t}} \int_{R^m} \lambda_s d\bar{r} d\tau \right] \int_{t_{N_t}}^{t_{N_t}+\Delta t} \int_{\bar{r}_{N_t} \epsilon c_{N_t}} \lambda_n d\bar{r} d\tau \exp\left[- \int_{t_{N_t-1}}^{t_{N_t}} \int_{R^m} \lambda_n d\bar{r} d\tau \right]}{\Delta t \prod_{j=1}^{m} \rho_{N_t j}} \qquad (63)$$

By repeated use of l'Hospital's rule and Leibnitz's rule, this reduces to

$$p(h_j^{N_t}(N_t) = 0, \bar{z}(N_t) | h_j^{N_t-1}, Z^{N_t-1})$$

$$= \lambda_n(t_{N_t}, \bar{r}_{N_t}) \exp\left[- \int_{t_{N_t-1}}^{t_{N_t}} \int_{R^m} \lambda d\bar{\pi} d\tau \right] \tag{64}$$

where

$$\lambda = \lambda(t, \bar{r}, \bar{x}(t)) = \lambda_s(t, \bar{r}, \bar{x}(t)) + \lambda_n(t, \bar{r}) \tag{65}$$

By the same argument, it can be shown that

$$p(h_j^{N_t}(N_t) = 1, \bar{z}(N_t) | h_j^{N_t-1}, Z^{N_t-1})$$

$$= \lambda_s(t_{N_t}, \bar{r}_{N_t}, \bar{x}(t_{N_t}))\exp\left[-\int_{t_{N_t-1}}^{t_{N_t}} \int_{R^m} \lambda d\bar{r} d\tau\right] \qquad (66)$$

We now have all the pieces necessary to calculate the weighting factors in equation (36). By substituting equation (64) or (66) as appropriate into equation (57), it can be seen that

$$\Pr[h_j^{N_t} | Z^{N_t}] = \begin{cases} \dfrac{\lambda_n(t_{N_t}, \bar{r}_{N_t})}{\lambda(t_{N_t}, \bar{r}_{N_t}, \bar{x}(t_{N_t}))} \cdot \Pr[h_j^{N_t-1} | Z^{N_t-1}] \\ \qquad \text{if } h_j^{N_j}(N_t) = 0 \\ \dfrac{\lambda_s(t_{N_t}, \bar{r}_{N_t}, \bar{x}(t_{N_t}))}{\lambda(t_{N_t}, \bar{r}_{N_t}, \bar{x}(t_{N_t}))} \cdot \Pr[h_j^{N_t-1} | Z^{N_t-1}] \\ \qquad \text{if } h_j^{N_t}(N_t) = 1 \end{cases} \qquad (67)$$

where the denominator is given in equation (65).

Thus, equation (36) defines our multiple model adaptive estimator, where the weighting factors are defined by equation (67) and $\hat{\underset{\sim}{x}}_j(t)$ is the Snyder and Fishman filter estimate of $\underset{\sim}{\bar{x}}(t)$ given the j^{th} hypothesis. This development is for the case of a known $\bar{x}(t)$, as assumed just prior to equation (42). The

$$Pr[\cdot] = \frac{\lambda_{s \text{ or } n}}{\lambda_s + \lambda_n} \cdot Pr[\cdot] \tag{68}$$

form provides insight into the nature of the weighting process and a similar form will be seen in the estimator developed in Chapter IV.

It is trivial, however, to estimate $\bar{x}(t)$ when we have assumed that it is a known deterministic function. We could correct this by allowing $\underset{\sim}{\bar{x}}(t)$ to be random and proceed as before to solve for the weighting factors in equation (41). The denominator of equation (41), the sample function density of equation (50), is actually $p(Z^{N_t} | \underset{\sim}{\bar{x}}(t)=\underset{\sim}{\bar{\xi}}(t))$ in this case and we can write

$$p(Z^{N_t}) = \int_{R^n} p(Z^{N_t} | \underset{\sim}{\bar{x}}(t) = \bar{\xi}(t)) p(\bar{\xi}(t)) d\bar{\xi} \qquad (69)$$

where $\bar{\xi}(t)$ is the dummy variable for $\bar{x}(t)$. Because $\underset{\sim}{\bar{x}}(t)$ is defined as the output of a linear Gaussian system, $p(\bar{x}(t))$ exists and the integration in equation (69) can, in principle, be performed although it is complicated and a closed form solution might not be possible.

Similarly, the numerator of equation (41) must be developed given $\underset{\sim}{\bar{x}}(t)$ as random and then averaging over the statistics of the $\underset{\sim}{\bar{x}}(t)$ process.

A further complication arises when we add feedback to achieve some sort of optimal control for the system. One form of the feedback could be as a control input $\bar{c}'(t, Z^{N_t})$ to the equation

$$d\underset{\sim}{\bar{x}}(t) = \underline{F}(t)\underset{\sim}{\bar{x}}(t)\, dt + \underline{G}(t) d\underset{\sim}{\bar{u}}(t) + \bar{c}'(t, Z^{N_t}) \qquad (70)$$

where $\bar{c}'(t, Z^{N_t})$ is an n dimensional control vector generated in some optimal manner from the observed process.

The presence of $\bar{c}'(t,Z^{N_t})$ will affect the form of the density $p(\bar{x}(t))$ and may make the integrations even more complicated, if not intractable.

The serious computational problems which arise with random $\underset{\sim}{\bar{x}}(t)$ and feedback control motivate us to consider modeling concepts other than the probability density approach taken in this chapter.

II.5 Summary

In this chapter, the physical model of the observed process is presented in which a space-time point process signal is observed, corrupted by space-time point process noise. The processes are assumed to be conditionally Poisson and statistically independent of each other. General MMAE techniques are discussed and the equations for MMAE on this point process problem are developed for non random rate parameters λ_s and λ_n. The development is based on probability density functions. The form of this estimator will provide insight in later chapters. The difficulties associated with using a probability density function approach for random rate parameters are discussed and motivation for the measure theory approach of Chapter III is presented.

In Chapter III, the observed physical process described in this chapter is modeled as a doubly stochastic space-time Poisson point process. This "cross product space" model and a measure theory approach to the statistics will allow us to overcome both the complex integration and feedback problems encountered in this chapter.

III. Cross Product Space Model

III.1 Introduction

The difficulties encountered in Chapter II in obtaining useful evaluations of the equations for multiple model adaptive estimation stem from the basic approach taken. The uncertainties were modeled in terms of probability density functions of random processes. This results in expressions which are very difficult to evaluate (equation (69) for example) except for trivial cases such as a known, non-random $\bar{x}(t)$. The addition of feedback control further complicates evaluation of the estimator equations.

In this chapter, a fundamentally different modeling approach is taken, but one which coincides well with the physical problem under consideration and which will allow us to make use of the MMAE development in Chapter II. The observed process is modeled as a doubly stochastic space-time point process on $[t_0,t)XR^m$. The statistics are defined using measure theory concepts on a cross product of two probability spaces and feedback control is included in the basic model. Some necessary results from Fishman, Reference 11, are presented in Sections III.2 and III.3, including a definition of regularity for a doubly stochastic

space-time point process. The implications of regularity are discussed in Section III.4 This section presents the tools necessary for overcoming the difficulties encountered in Chapter II in deriving the weighting factors for the multiple model adaptive estimator. The main result is Theorem III.6, the representation theorem, which provides a method for calculating a posteriori probabilities of the form needed for the weighting factors.

In Section III.4, an analytic description of the beam point process estimation problem is given in terms of the cross product space model and a regularity proof is given for this analytical form.

III.2 Doubly Stochastic Space-Time Point Process

In general terms, a doubly stochastic space-time point process is a space-time point process in which some parameter of the process is itself random. Our physical model for the beam problem fits this description: we observe a space-time point process on $[t_0,t)XR^m$ and we wish to estimate the state of the Markov process which determines the location of the centroid of the observed process. We have additionally assumed that, given $\bar{x}(t)$ and $\lambda_n(\cdot)$, the process is Poisson. This assumption is not necessary, however, for the results presented in this section. As defined in equation (4), $\bar{x}(t)$ is random and we wish to allow other terms in $\lambda_s(\cdot)$ to be random. A random

noise rate parameter is also desired and can be used to model uncertainties in external noise sources or in the detector. Thus, the particle beam problem can be described readily in terms of a doubly stochastic point process.

We begin a more thorough description by defining a probability space (Ω_s, A_s, P_s) where Ω_s is a nonempty set, is a Borel field (sigma field) of subsets of Ω_s and P_s is a probability measure on A_s. This probability space corresponds to events we cannot observe directly. For all $\omega_s \varepsilon \Omega_s$ let there be a probability space $(\Omega, \mathcal{B}, P(\cdot;\omega_s))$ where the probability measure $P(\cdot;\omega_s)$ is dependent on the event in Ω_s, $\mathcal{B}$ is a Borel field of subsets of Ω and $P(\cdot;\omega_s)$ is a probability measure on $\mathcal{B}$. If for every $B \varepsilon \mathcal{B}$, $P(B;\cdot)$ is measurable on (Ω_s, A_s) then it can be shown (Ref. 32) that a unique joint probability measure P' exists on $(\Omega_s \times \Omega, A_s \times \mathcal{B})$ (where $A_s \times \mathcal{B}$ denotes the product algebra of A_s and $\mathcal{B}$ (Ref. 32:71)) such that

$$P'(A \times B) = \int_A P(B;\omega_s)\, P_s(d\omega_s) \qquad (71)$$

$$A \varepsilon \mathcal{B}_s$$

$$B \varepsilon \mathcal{B}$$

and a unique probability measure $\bar{P}$ exists on $(\Omega,\mathcal{B})$

$$\bar{P}(B) = P'(\Omega_S \times B) = \int_{\Omega_S} P(B;\omega_S) P_S(d\omega_S) \qquad (72)$$

$$B \in \mathcal{B}$$

In addition, if $W(\omega_S;\omega)$ is a random variable on $(\Omega_S \times \Omega, A_S \otimes \mathcal{B})$ then

$$\int_{\Omega_S \times \Omega} W dP' = \int_{\Omega_S} \left[\int_{\Omega} W(\omega_S;\omega) P(d\omega;\omega_S) \right] P_S(d\omega_S) \qquad (73)$$

Let U and V be mappings from $[t_0,\infty) \times \Omega$ to $[t_0,\infty) \times R^m$

$$U:[t_0,\infty) \times \Omega \to [t_0,\infty) \times R^m$$

$$V:[t_0,\infty) \times \Omega \to [t_0,\infty) \times R^m \tag{74}$$

in which an event $B \varepsilon \mathcal{B}$ is mapped into a set of space time points in $[t_0,\infty) \times R^m$. The underlying probability space for the process U is $(\Omega,\mathcal{B},P(\cdot;\omega_s))$. The underlying probability space for the process V is $(\Omega,\mathcal{B},\bar{P})$. Note that U and V have identical pre-image and image spaces. The distinction between the two processes is that we explicitly show the dependence on ω_s for process U .

This definition of a doubly stochastic space-time point process as a mapping from a cross product of probability spaces gives us a convenient framework for describing the particle beam problem. We observe point events, $\bar{z}(i)$, $i \varepsilon 1,2,\ldots,N_t$ which are generated by a conditionally Poisson process. The rate parameter, $\lambda(\cdot)$, of the Poisson process is defined by equation (8) and is itself a function of the random process $\underset{\sim}{\bar{x}}(t)$ and $\lambda_n(\cdot)$. If we let the probability space (Ω_s,A_s,P_s) specify the random nature of

$\underset{\sim}{\bar{x}}(t)$ and (the potentially random) $\lambda_n(\cdot)$, then given ω_s , the process U is Poisson. Furthermore, the doubly stochastic space-time point process V is defined for this application. Figure 5 summarizes this modeling concept.

III.3 Regularity

Several important results can be obtained if the doubly stochastic space-time point process (modeled as a mapping from a cross product of probability spaces) is regular. We first consider the definition of regularity and then some useful implications of regularity. Some preliminary definitions and notation are necessary prior to defining regularity.

Definition III-1. Let t be an element of $[t_0,T)$ and let N_t be defined as before. Let $\mathcal{B}_t$ denote the subsigma field of $\mathcal{B}$ generated by the random variables

$$\{N_t,(t_1,\bar{r}_1), \ldots ,(t_{N_t},\bar{r}_{N_t})\} \qquad \blacksquare$$

This can be interpreted as the sigma field generated by the sample paths of the point process up to time t.

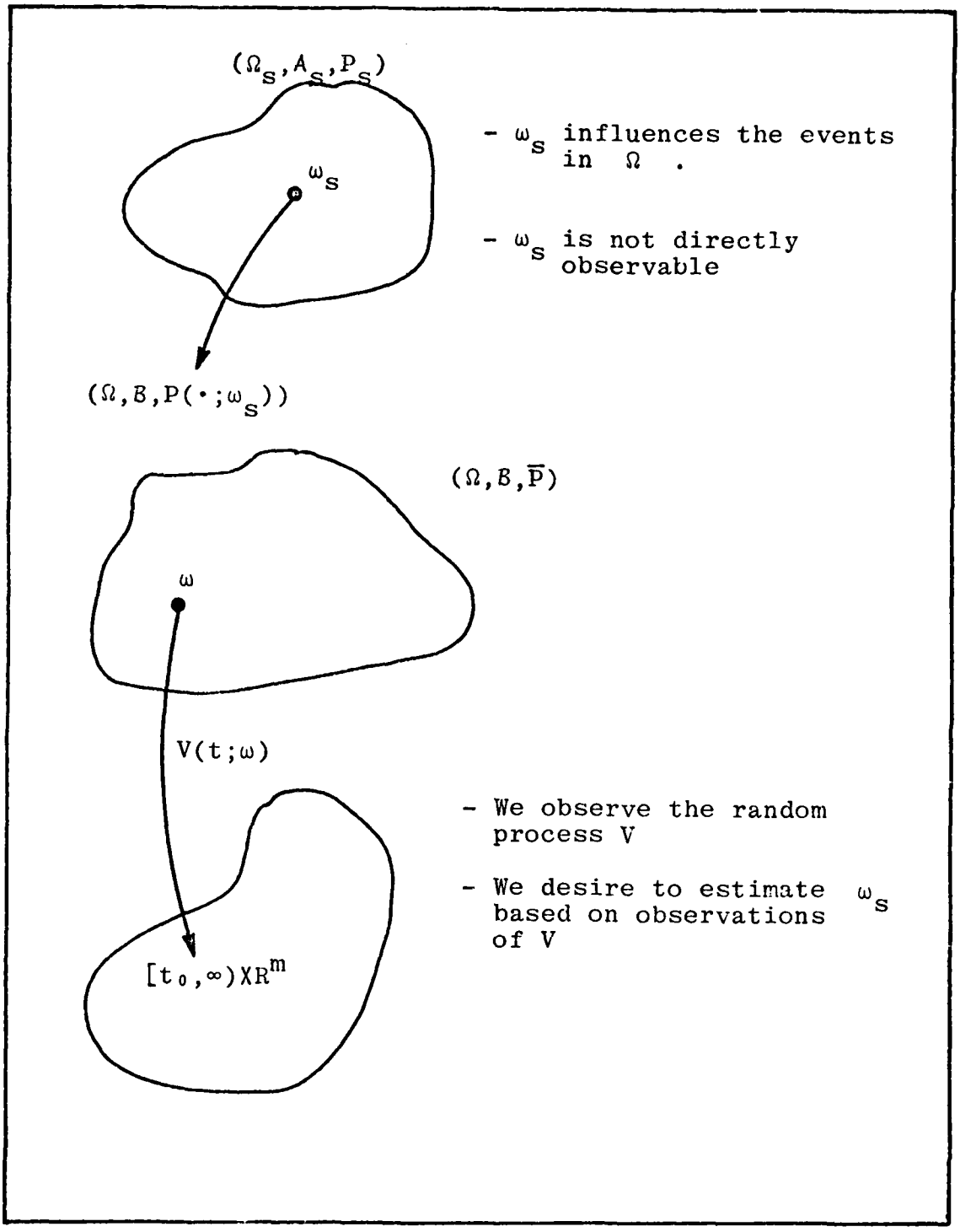


Figure 5. The Spaces

Definition III-2. A space-time point process is called conditionally orderly if for each point $(t,\bar{r})\varepsilon[t_0,T)X\ Y$

$$\lim_{\substack{\Delta t\to 0\\ \bar{\rho}\to 0}} \frac{\Pr[N([t,t+\Delta t)Xc(\bar{r},\bar{\rho}))\geq 2 \mid \mathcal{B}_t](\omega)}{\Pr[N([t,t+\Delta t)Xc(\bar{r},\bar{\rho}))\geq 1 \mid \mathcal{B}_t](\omega)} = 0 \qquad (75)$$

w.p.1

whenever the denominator does not vanish. ■

Conditional orderliness essentially guarantees that the observed events are distinct; the probability of two or more point events occurring at the same time and spatial location is zero.

Definition III-3. Let $\bar{z}(i)\varepsilon[t_0,t)\ X\ Y$, $i\varepsilon$ $1,2,3,\ \ldots\ ,N_t$, be defined as in equation (38) and let $\underset{\sim}{\bar{z}}(i)$ denote a random variable for the observation at time t_i and spatial location $\bar{r}_i$

$$\underset{\sim}{\bar{z}}(i) = \begin{bmatrix} \underset{\sim}{t}_i \\ - - - \\ \underset{\sim}{\bar{r}}_i \end{bmatrix} \tag{76}$$

The j^{th} order distribution function is defined for $j \geq 1$ by

$$F_j(Z^j) \triangleq F_j(\bar{z}(1),\bar{z}(2), \ldots ,\bar{z}(j))$$

$$= Pr[\underset{\sim}{\bar{z}}(1) \leq \bar{z}(1), \ldots ,\underset{\sim}{\bar{z}}(j) \leq \bar{z}(j)]$$

$$\bar{z}(i) \epsilon [t_0,t) \times Y \tag{77}$$

$$i = 1,2, \ldots ,j$$

where the vector inequality is taken as an inequality on each corresponding element of the two vectors. ■

We can now define the conditions a space-time point process must satisfy in order to be regular.

Definition III-4. A space-time point process which maps into $[t_0,t)XY$ is regular if the following four conditions are satisfied:

(a) Each distribution function

$$F_j(Z^j) \triangleq F_j\ (\bar{z}(1), \ \ldots\ , \bar{z}(j)) \tag{78}$$

$$j = 1,2, \ \ldots\ , N_t$$

is absolutely continuous on

$$R^{(m+1)j} \tag{79}$$

(b)

$$Pr[(t_{j+1}, \bar{r}_{j+1}) \varepsilon (t_j, t)XY | \mathcal{B}_{t_j}] < 1 \tag{80}$$

w.p.1

for all finite t , j=0,1,2, ...

(c) The point process is conditionally orderly. (81)

(d)

$$\Pr[N([t_0,t)XY)<\infty] = 1 \tag{82}$$

for all finite t (Ref. 11:79) ■

Condition (b) requires that the conditional probability of a new point occurring anywhere in Y, for finite t, is less than one; that is, there are no guaranteed points. Condition (d) requires that the number of points in Y is finite, for finite t.

Our processes will be modeled as doubly stochastic point processes. Regularity for this case is defined as follows.

Definition III-5. A regular doubly stochastic space-time point process $V:[t_0,T)X\Omega\rightarrow[t_0,T)XY$ is a doubly stochastic point process such that U (equation 74) is a regular space-time point process for each $\omega_s \varepsilon \Omega_s$ (Ref. 11:105). ■

Definition III-5 provides the conditions under which the doubly stochastic space-time point process V is regular. In order to use the results of the next section,

we require that V be a regular space-time point process (versus a regular doubly stochastic space-time point process). The distinction is that for a regular doubly stochastic space-time point process, V , we must specify ω_s to insure the regularity of V . If V is a regular space-time point process, we do not need to specify ω_s to use the results of regularity (even though V is dependent on ω_s). The necessary conditions for this are provided by Theorem III-1.

The $\phi(\cdot)$ functions in the following theorem are termed hazard functions and they specify the infinitesimal properties of regular space-time point processes. A hazard function is the conditional instantaneous rate of occurrence of new events per space-time volume. The hazard function, $\phi(\cdot)$, for the general regular space-time point process corresponds exactly to the rate parameter, $\lambda(\cdot)$, for a Poisson space-time point process. The existence and usefulness of hazard functions are discussed in Section III.4.1. For a proof of Theorem III.1, see reference 11 pages 106-116.

Theorem III-1. Let V be a regular doubly stochastic space-time point process such that

(a) If $B \in \mathcal{B}$ and $t \in [t_0, T)$ then $P(B; \omega_s | \mathcal{B}_t)(\omega)$ is measurable with respect to the product sigma field

$$A_s \otimes \mathcal{B} \tag{83}$$

(b) For each point $(t, \bar{r}) \in [t_0, T) \times Y$

$$\phi(t, \bar{r}) = E\{\phi(t, \bar{r}; \omega; \omega_s)\}$$

$$= \int_{\Omega \times \Omega_s} \phi(t, \bar{r}; \omega; \omega_s) P'(d\omega \times d\omega_s) < \infty \tag{84}$$

(c) $\phi(t, \bar{r}; \omega; \omega_s)$ is measurable with respect to

$$[t_0, t) \times R^m \times \mathcal{B} \times A_s \tag{85}$$

Then V is a regular space-time point process.

■

III.4 Implications of Regularity

When a space-time point process satisfies the conditions for regularity, it can be shown that the associated hazard function for the process exists (Ref. 11:83-89). A sample function density can be written in terms of the hazard function in several illustrative forms. Of direct importance to the estimation and control problem for a point process signal in point process noise is that the hazard function for the doubly stochastic space-time point process V is dependent on ω_s and ω. The ω dependence allows for feedback control of the system from the observations in $[t_0,T)XR^m$. Regularity also provides a means of calculating $P_s(A|\mathcal{B}_t)$, the probability of an event $A\epsilon A_s$ given the sub-sigma algebra generated by the measurements, $\mathcal{B}_t$. If we model, in Ω_s, the uncertainty of whether the noise process or signal process caused an observed point event, then we can use $P_s(A|\mathcal{B}_t)$ to derive the probability that a particular hypothesis sequence (represented by A) occurred, given $\mathcal{B}_t$. This is the key result necessary to develop the individual filter weights for the multiple model adaptive estimator.

As a result of regularity, we can also develop a direct estimator (as opposed to a multiple model adaptive estimator) for the process $\underset{\sim}{\bar{x}}(t)$ when the observations are corrupted by point process noise. The direct estimator is

developed in Appendix A. The multiple model adaptive estimator approach results in much simpler individual expressions to evaluate than does the direct estimator. The tradeoff is the growing requirement for calculation and memory necessary for the MMAE approach.

The specific implications of regularity are presented in the rest of this section. They are due to Fishman (Ref. 11) and are presented without proof as background for the multiple model adaptive estimation filter development in Chapter IV.

III.4.1 Hazard Functions. We let W be a space-time point process, $W:[t_0,T)X\Omega \rightarrow [t_0,t)XR^m$. If this space-time point process is regular, that is, it satisfies definition III-4, then the hazard function, $\phi(t,\bar{r};\omega),\omega\varepsilon\Omega$ exists (Ref. 11:83-89). The infinitesimal properties of the regular space-time point process are described in the following theorem.

Theorem III-2. The following limit holds (w.p.1) for almost all $(t,\bar{r})\varepsilon[t_0,T)XY$

$$\lim_{\substack{\Delta t \downarrow 0 \\ \bar{\rho} \to 0}} (\Delta t \bar{\rho})^{-1} \Pr[N([t,t+\Delta t) \times c(\bar{r},\bar{\rho})) \geq 1 | \mathcal{B}_t](\omega)$$

$$= \lim_{\substack{\Delta t \downarrow 0 \\ \bar{\rho} \to 0}} (\Delta t \bar{\rho})^{-1} \Pr[N([t,t+\Delta t) \times c(\bar{r},\bar{\rho})) = 1 | \mathcal{B}_t](\omega) \qquad (86)$$

$$= \phi(t,\bar{r};\omega)$$

Proof: (Ref 11:89) ∎

If a doubly stochastic space-time point process V is regular (Definition III-5), then each process U is regular for a given $\omega_s \varepsilon \Omega_s$. A hazard function exists for each doubly stochastic regular space-time point process V and is of the form $(t,\bar{r};\omega;\omega_s)$, $\omega \varepsilon \Omega, \omega_s \varepsilon \Omega_s$. The dependence on ω_s is due to the requirement that U is regular given ω_s.

If the doubly stochastic regular space-time point process V satisfies Theorem III-1, then V is a regular

space-time point process (note, the "doubly stochastic" qualifier has been dropped). Because the space-time point process V is regular, a hazard function exists in the form $\overline{\phi}(t,\overline{r};\omega)$. The evaluation of $\overline{\phi}(t,\overline{r};\omega)$ is presented in the following theorem.

Theorem III-3. Let V be a doubly stochastic space-time point process satisfying Theorem III-1. Then (w.p.1) the following equation is valid for almost all $(t,\overline{r})\varepsilon[t_0,T) \times Y$

$$\phi(t,\overline{r};\omega) = \hat{\phi}(t,\overline{r};\omega) \triangleq E_s\{\phi(t,\overline{r};\omega;\omega_s)|A_0\otimes\mathcal{B}_t\} \tag{87}$$

where $A_0\varepsilon A_s$ is the set $\{\emptyset,\Omega_s\}$, and $\emptyset$ is the empty set.

Proof: (Ref. 11:117) ∎

The conditioning on the sub-sigma field $A_0\otimes\mathcal{B}_t$ is

equivalent to conditioning on the "measurement history" $\mathcal{B}_t$ and no further information about Ω_s .

Sample function densities for regular space-time point processes can be written in terms of hazard functions, as shown in the following theorem.

Theorem III-4. Let V be a regular space-time point process which maps into $[t_0,T) \times Y$. Let $t\in[t_0,T)$ and denote the sample function density on $[t_0,t) \times Y$ as $L_t(\omega)$. Then (w.p.1)

$$L_t(\omega) \overset{a.s.}{=} \exp\left[-\int_{t_0}^{t}\int_{Y}\phi(\tau,\bar{r};\omega)d\bar{r}d\tau\right] \tag{88}$$

for $N_t = 0$

$$L_t(\omega) \overset{a.s.}{=} \prod_{i=1}^{N_t}\phi(t_i,\bar{r}_i;\omega)\exp\left[-\int_{t_0}^{t}\int_{Y}\phi(\tau,\bar{r};\omega)d\bar{r}d\tau\right] \tag{89}$$

$N_t \geq 1$

Proof: (Ref 11:92) ■

Equation (89) can be written alternately as

$$L_t(\omega) = \exp\left[-\int_{t_0}^{t}\int_{Y} \phi(\tau,\bar{r};\omega)d\bar{r}d\tau + \sum_{i=1}^{N_t} \ln\phi(t_i,\bar{r}_i;\omega) \right] \qquad (90)$$

or

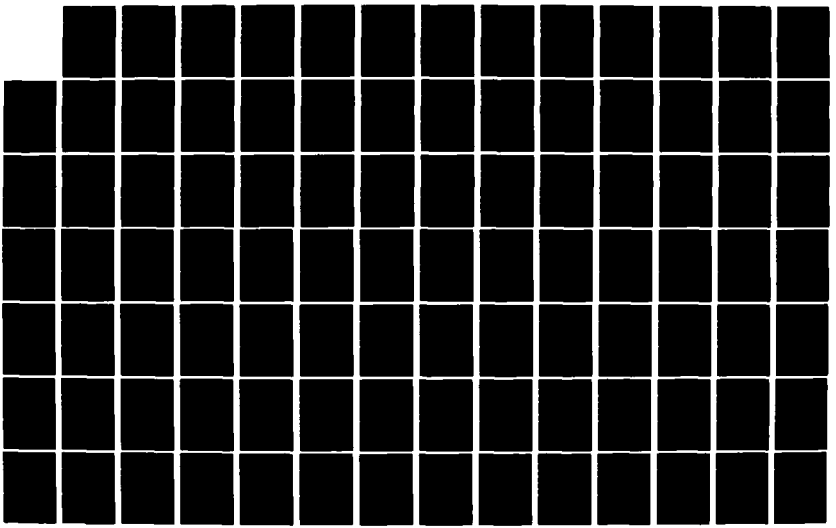
$$L_t(\omega) = \exp\left[-\int_{t_0}^{t}\int_{Y} \phi(\tau,\bar{r};\omega)d\bar{r}d\tau + \int_{t_0}^{t}\int_{Y} \ln\phi(\tau,\bar{r};\omega)N(d\tau X d\bar{r}) \right] \qquad (91)$$

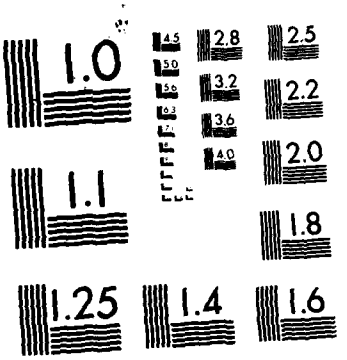
where the last integral in equation (91) is called a stochastic counting integral and is defined by

$$\int_{t_0}^{t}\int_{Y} \ln\phi(\tau,\bar{r};\omega)N(d\tau X d\bar{r}) \triangleq \sum_{i=1}^{N_t} \ln\phi(t_i,\bar{r}_i;\omega) \qquad (92)$$

AD-A124 827 MULTIPLE MODEL ADAPTIVE ESTIMATION FOR SPACE-TIME POINT PROCESS OBSERVATIONS(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF ENGINEERING D E MEER 2/3

UNCLASSIFIED AUG 82 AFIT/DS/EE/82-2 F/G 12/1 NL





We can similarly define the sample function density of the doubly stochastic regular space-time point process V as

$$L_t(\omega;\omega_s) = \prod_{i=1}^{N_t} \phi(t_i,\bar{r}_i;\omega;\omega_s)\exp\left[-\int_{t_0}^{t}\int_{Y} \phi(\tau,\bar{\pi};\omega;\omega_s)d\bar{\pi}d\tau\right] \quad (93)$$

$$N_t \geq 1$$

and the $N_t=0$ case is exactly analogous to equation (88).

Furthermore, we can write the sample function density of the regular space-time point process V in terms of $\hat{\phi}(\cdot)$ (recall the definition in equation (87)).

Theorem III-5. Let the doubly stochastic regular space-time point process V satisfy Theorem III-1 (V is regular). Then

$$\bar{L}_t(\omega) = \prod_{i=1}^{N_t} \hat{\phi}(t_i,\bar{r}_i;\omega)\exp\left[-\int_{t_0}^{t}\int_{Y}\hat{\phi}(\tau,\bar{r};\omega)d\bar{r}d\tau\right] \tag{94}$$

$$= \int_{\Omega_S} L_t(\omega;\omega_s)P_s(d\omega_s)$$

$$= E_s\{L_t(\omega;\omega_s)\}$$

Proof: (Ref. 11:118-119) ■

With these preliminaries, we can now write the representation theorem. This gives us the means of evaluating the probability of an event $A\varepsilon A_s$ (which is not directly observable) given $\mathcal{B}_t$.

Theorem III-6, Representation Theorem. Let a doubly stochastic space-time point process, V satisfy the conditions of Theorem III-1 on $[t_0,T)XY$. Then

$$P_s(A|\mathcal{B}_t) = \frac{\int_A L_t(\omega;\omega_s)P_s(d\omega_s)}{\int_{\Omega_s} L_t(\omega;\omega_s)P_s(d\omega_s)} \tag{95}$$

$$= \frac{E_s\{L_t(\omega;\omega_s)|\omega_s \varepsilon A\}P_s(A)}{E_s\{L_t(\omega;\omega_s)\}} \tag{96}$$

$A \varepsilon A_s$

Proof: (Ref. 11:127) ■

As mentioned previously, it is the representation theorem which will allow us to evaluate the $Pr[h_j^{N_t}$ is correct $|\mathcal{B}_t]$ terms in the multiple model adaptive estimator posed in Chapter II.

With the tools of this section, the next steps are to:

(a) Develop an analytical description of the particle beam problem which fits the cross product space modeling concept.

(b) Prove that the process U specified by the analytical model is regular (Definition III-4).

(c) Prove that the space-time point process V specified by the analytical model is regular (Theorem III-1).

This is accomplished in the next section.

III.5 Analytical Cross Product Space Model

The conceptual cross product space model has been presented already in Figure 5. In this section, we develop an analytical description for the particle beam problem which fits the cross product space concept. The analytical description is necessary for evaluating the multiple model adaptive estimator in Chapter IV.

III.5.1 The Observed Space, Ω. Let U be a random point process $U:[t_0,T)X\Omega \rightarrow [t_0,T)XR^m$. We assume that the process is Poisson, conditioned on $\omega_s \varepsilon \Omega_s$, where Ω_s is described in Section III.5.2. We let the rate parameter for

the conditional Poisson process be $\lambda(t,\bar{r};\omega;\omega_s)$ where $\omega\varepsilon\Omega$, $\omega_s\varepsilon\Omega_s$. Note that $\lambda(\cdot)$ is random; however, the under tilde notation has been dropped because the dependence on ω and ω_s is shown explicitly.

The basic assumption made here is that the process is Poisson, conditioned on ω_s . This assumption is made because a conditionally Poisson process models the photoelectron events adequately (Ref. 13:49-55). As will be shown later in this section, if we restrict $\lambda(\cdot)$ so that

$$\int_{t_0}^{t}\int_{Y}\lambda(\tau,\bar{\pi};\omega;\omega_s)d\bar{\pi}d\tau < \infty \qquad (97)$$

for all $t < \infty$

then U is a doubly stochastic regular space-time point process, the hazard function

$$\phi(t,\bar{r};\omega;\omega_s) = \lambda(t,\bar{r};\omega;\omega_s) \qquad (98)$$

exists, and the dependence of $\lambda(\cdot)$ on both ω and ω_s is justified.

III.5.2 The Unobserved Space Ω_S. Let (Ω_S, A_S, P_S) be a probability space in which $\omega_S \varepsilon \Omega_S$. Further, let it be a cross product of three distinct probability spaces

$$(\Omega_{S_1}, A_{S_1}, P_{S_1})$$

$$(\Omega_{S_2}, A_{S_2}, P_{S_2}) \tag{99}$$

$$(\Omega_{S_3}, A_{S_3}, P_{S_3})$$

(which correspond to the randomness in the signal, noise, and hypothesis sequences, respectively) such that the non-empty set Ω_S is defined as

$$\Omega_S \triangleq \Omega_{S_1} \times \Omega_{S_2} \times \Omega_{S_3} \tag{100}$$

and the sigma field A_S is defined as

$$A_S \triangleq A_{S_1} \otimes A_{S_2} \otimes A_{S_3} \tag{101}$$

The sigma field A_{S_1} is composed of subsets of Ω_{S_1}, and P_{S_1} is a probability measure on A_{S_1}. The terms A_{S_2}, A_{S_3}, P_{S_2}, and P_{S_3} are similarly defined for the corresponding probability spaces. Thus, ω_S is specified completely by $\omega_{S_1} \varepsilon \Omega_{S_1}$, $\omega_{S_2} \varepsilon \Omega_{S_2}$, and $\omega_{S_3} \varepsilon \Omega_{S_3}$.

Let the rate parameter for the observed conditionally Poisson process be

$$\lambda(t,\bar{r};\omega;\omega_S) = \alpha_1(t,\bar{r};\omega_{S_3})\lambda_S(t,\bar{r};\omega;\omega_{S_1}) + \alpha_0(t,\bar{r};\omega_{S_3})\lambda_n(t,\bar{r};\omega;\omega_{S_2}) \tag{102}$$

The signal rate parameter, $\lambda_S(t,\bar{r};\omega;\omega_{S_1})$, is defined as

$$\lambda_s(t,\bar{r};\omega;\omega_{s_1})$$

$$= \Lambda(t)\exp\left\{-\tfrac{1}{2}[\bar{r}-\underline{H}(t)\bar{x}(t;\omega_{s_1})+\bar{c}(t;\omega)]^T \underline{R}^{-1}(t)\right.$$

$$\left.\cdot[\bar{r}-\underline{H}(t)\bar{x}(t;\omega_{s_1})+\bar{c}(t;\omega)]\right\} \qquad (103)$$

where $\bar{x}(t;\omega_{s_1})$ is an n dimensional random variable with domain in the probability space $(\Omega_{s_1},A_{s_1},P_{s_1})$, $\bar{c}(t;\omega)$ is a (possible) feedback control, and all other terms in equation (103) are the same as in equation (3). The process $x(t;\omega_{s_1})$, is defined as the solution to

$$d\bar{x}(t;\omega_{s_1}) = \underline{F}(t)\bar{x}(t;\omega_{s_1})dt+\underline{G}(t)d\bar{u}(t;\omega_{s_1})+\bar{c}'(r;\omega) \qquad (104)$$

$$\bar{x}(t_0;\omega_{s_1}) = \underset{\sim}{\bar{x}}_0$$

$$t_0 \leq t$$

where $\bar{c}'(t;\omega)$ is a (possible) control input.

The difference between equations (3) and (4) and equations (103) and (104) is that in the latter we now specify the randomness in $\lambda_s(t,\bar{r};\omega;\omega_{s_1})$ (other than a possible control feedback) via the probability space $(\Omega_{s_1},A_{s_1},P_{s_1})$.

There are two significant points to be noted for future expansion of the model. First, by using this cross product space model, we could let other components of $\lambda_s(t,\bar{r};\omega;\omega_{s_1})$ be dependent on ω_{s_1}. For example, we could model an uncertain rate parameter "amplitude" as $\Lambda(t;\omega_{s_1})$ in equation (103) where the probability space $(\Omega_{s_1},A_{s_1},P_{s_1})$ is expanded appropriately to specify both $\Lambda(t;\omega_{s_1})$ and $\bar{x}(t;\omega_{s_1})$. Second, because $\lambda(t,\bar{r};\omega;\omega_{s_1})$ is allowed to be a function of ω, we can include feedback control in this model. One method of doing this could be to include an additive control of the form $\bar{c}(t;\omega)$ as is shown in equation (103). In a beam tracking application, in which we are limited to a finite sized array of photodetectors, the control $\bar{c}(t;\omega)$ might be used to adjust the pointing of the detector to keep the current estimate of the projected position (estimate of the terms $\underline{H}(t)\bar{x}(t;\omega_{s_1})$) in the center of the array. If this were not done, the "signal" (maximum intensity point of the Gaussian shaped intensity profile) could conceivably "walk" off the detector array and no further useful information

would be observed from the signal process.

An alternate (or additional) method of control is provided by the term $\bar{c}'(t;\omega)$ in equation (104). In a beam pointing application, we can model the physical influence on the beam position as a control input $\bar{c}'(t;\omega)$ to the process $d\bar{x}(t;\omega_{s_1})$ where $\bar{x}$ is a relative rather than an absolute position.

The noise rate parameter is defined as a scalar random variable

$$\lambda_n(t,\bar{r};\omega;\omega_s) = \lambda_n(t,\bar{r};\omega;\omega_{s_2}) \tag{105}$$

with domain specified by the probability space $(\Omega_{s_2}, A_{s_2}, P_{s_2})$. Note that dependence on (feedback control) is allowed in this model. As noted before, the regularity proof requires that:

$$\int_{t_0}^{t}\int_{Y}\lambda_n(\tau,\bar{\pi};\omega;\omega_{s_2})d\bar{\pi}d\tau < \infty \tag{106}$$

$$E\{\lambda_n(t,\bar{r},\omega;\omega_{s_2})\}<\infty$$

$$t<\infty$$

$$Y \subseteq R^m$$

The probability space $(\Omega_{s_3}, A_{s_3}, P_{s_3})$ and the coefficients $\alpha_0(t,\bar{r};\omega_{s_3})$ and $\alpha_1(t,\bar{r};\omega_{s_3})$ allow us to describe analytically the conditioning in expressions of the form (for example)

$$E_s\{(\cdot)|A\}P_s(A) = \int_A (\cdot)P_s(d\omega_s) \tag{107}$$

where $(\cdot)$ is an event of interest and $A \epsilon A_s$ is the event associated with one of the 2^{N_t} possible sequences $h_j^{N_t}$. Evaluation of equations of the form shown above are necessary in Chapter IV to derive the multiple model adaptive estimator equations.

Let $(\Omega_{s_3}, A_{s_3}, P_{s_3})$ be a discrete probability space where each $\omega_{s_3} \epsilon \Omega_{s_3}$ is associated with one of the hypothesis sequences. Note that at time t, there are 2^{N_t} possible sequences, $h_j^{N_t}$, $j \epsilon 1,2, \ldots 2^{N_t}$, and that the number of possible events in Ω_{s_3} doubles as each new measurement is observed.

The α coefficients are defined as

$$\alpha_0(t,\bar{r};\omega_{s_3}) = \begin{cases} 0 & \text{At measurement times and locations when it is given (or assumed) that the event at time } t \text{ was due to signal.} \\ 1 & \text{Otherwise.} \end{cases} \quad (108)$$

$$\alpha_1(t,\bar{r};\omega_{s_3}) = \begin{cases} 0 & \text{At measurement times and locations when it is given (or assumed) that the event at time } t \text{ was due to noise.} \\ 1 & \text{Otherwise.} \end{cases} \quad (109)$$

$$t_0 \leq t$$

and

$$\alpha_0(t,\bar{r};\omega_{s_3}) = \alpha_1(t,\bar{r};\omega_{s_3}) = 1 \quad (110)$$

If there is no assumption (or given information) concerning which process (noise or signal) caused the observed events.

An example will help clarify the notation. For observation on $[t_0,t)XY$, let $N_t = 3$, and $Z^3 = \{(t_1,\bar{r}_1),(t_2,\bar{r}_2),(t_3,\bar{r}_3)\}, t_0<t_1<t_2<t_3<t$. We desire to evaluate an equation of the same form as equation (107). The assumed hypothesis sequence associated with the event A is, for example,

$$h_6^3 = \{h_6^3(1),h_6^3(2),h_6^3(3)\}$$

$$= \{1,1,0\} \qquad (111)$$

$$= \{\text{signal at } t_1, \text{ signal at } t_2, \text{ noise at } t_3\}$$

Recall that the superscript on the hypothesis sequence is the number of space-time events that have been observed and the subscript is the index of sequence under consideration, out of a possible 2^{N_t} possible sequences. The associated space-time history of the α coefficients is as shown in Figure 6. Because $\alpha_0(t,\bar{r};\omega_{S_3}) = 0$ at $(t_1,\bar{r}_1)$ and $(t_2,\bar{r}_2)$, the observed events at those times could only have

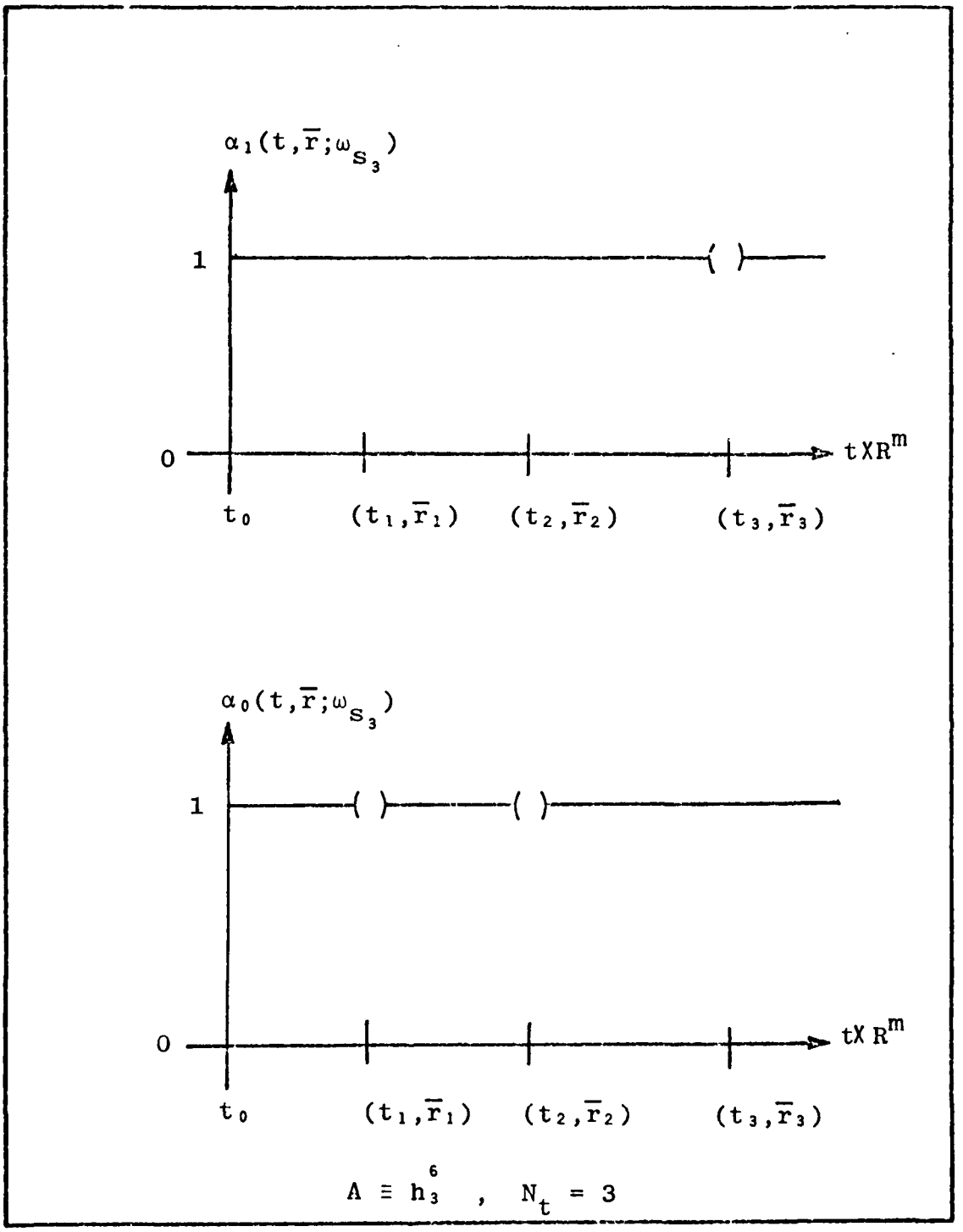


Figure 6. The α Coefficients

been caused by the signal process. Similarly, the event observed at $(t_3,\bar{r}_3)$ could only have been caused by the noise process. The α coefficients are specified as having a value of "1" between measurement times because, even though no point event occurred in these intervals for the observed realization, a point event could have occurred. The conditioning on $h_j^{N_t}$ specifies only what can or can not occur at the observed event times.

We now have an analytical method for modeling conditioning on some particular hypothesis sequence. In order to use Fishman's results, we must first prove that the space-time point process specified on the probability space $(\Omega,\mathcal{B},P(\cdot;\omega_s)$, and described analytically as above, is regular, conditioned on ω_s.

Theorem III-7. Let $(\Omega,\mathcal{B},P(\cdot;\omega_s))$ be a probability space and for each ω_s let U be a random point process $U:[t_0,T)X\Omega\rightarrow[t_0,T)XY$, which is Poisson conditioned on knowledge of $\omega_s\varepsilon\Omega_s$. Let the rate parameter of the conditionally Poisson process be $\lambda(t,\bar{r};\omega;\omega_s)$ as in equation (102) and let (Ω_s,A_s,P_s) be a probability space which is defined as a cross product of three probability spaces as described by equations (99), (100), and (101). If

$$\int_{t_0}^{t}\int_{Y}\lambda_n(t,\bar{r};\omega;\omega_{s_3})d\bar{r}d\tau < \infty \qquad (112)$$

$$t_0 \leq t < \infty$$

$$Y \subseteq R^m$$

then the space-time point process U is regular given $\omega_s \varepsilon \Omega_s$.

□

Proof. To prove Theorem III-7, we must show that the four conditions of Definition III-4 are satisfied for the analytical description given of $(\Omega_s, \mathcal{B}_s, P_s)$. Note that this theorem states that the point process U is regular given $\omega_s \varepsilon \Omega_s$, therefore in this proof, ω_{s_1} , ω_{s_2}, and ω_{s_3} are all known events in the analytical model of $\lambda(t,\bar{r};\omega;\omega_s)$.

(a) To show that: $F_j(Z^j)$ is absolutely continuous on $R^{(m+1)j}$:

By equation (77), the distribution function is defined as

$$F_j(Z^j) = \Pr[\underset{\sim}{\bar{z}}(1) \leq \bar{z}(1), \ldots, \underset{\sim}{\bar{z}}(j) \leq \bar{z}(j)] \tag{113}$$

We can write this in terms of a j^{th} order joint density function. The representation is formal in that we do not know if the density exists.

$$F_j(Z^j) = \int_{-\infty}^{\bar{z}(j)} \cdots \int_{-\infty}^{\bar{z}(1)} f(\bar{z}(1),\ldots,\bar{z}(j))d\bar{z}(1)\ldots d\bar{z}(j) \tag{114}$$

where the integrals are interpreted as

$$\int_{-\infty}^{\bar{z}(i)} \cdot d\bar{z}(i) = \int_{t_0}^{t_i} \int_{-\infty}^{h_{i_1}} \cdots \int_{-\infty}^{h_{i_m}} \cdot dh_{i_m} \ldots dh_{i_1} d\tau_i \tag{115}$$

if: $Y = R^m$

$i = 1,2,\ldots,j$

and the integral signs and differentials are nested from the innermost pair outward. If $Y \subset R^m$, the elemental integrals in equation (115) are over the regions

$$(\bar{r}_i \varepsilon Y) \cap (r_{i_k} \leq z_k(i)) \qquad (116)$$

$$k = 1,2,\ldots,m$$

The density in equation (114) is the probability of $\{\underset{\sim}{z}(1),\ldots,\underset{\sim}{z}(j)\}$ falling within some infinitesimal volume about Z^j. This is precisely the sample function density given in equation (50) and repeated here:

$$p(Z^j) = \prod_{i=1}^{j} \lambda(t_i,\bar{r}_i,\omega;\omega_s)\exp\left[-\int_{t_0}^{t}\int_{Y}\lambda(\tau,\bar{r};\omega;\omega_s)d\bar{r}d\tau\right] \qquad (117)$$

Since ω_s is given in the conditions for this theorem, the process U is conditionally Poisson and this expression for $f(Z^j)$ is valid. The sample function density exists. By the Radon-Nikodym Theorem (Ref. 54:34), if the density $f(Z^j)$ exists, then $F_j(Z^j)$ is absolutely continuous with respect to $R^{(m+1)j}$ and the condition is satisfied.

(b) To show that: $\Pr[(t_{i+1},\bar{r}_{i+1})\varepsilon(t_i,t)XY|\mathcal{B}_{t_i}] < 1$

w.p.1, $i = 0,1,....$

$$\Pr[(t_{i+1},\bar{r}_{i+1})\varepsilon(t_i,t)XY|\mathcal{B}_{t_i}] \tag{118}$$

$$= 1-\Pr[(t_{i+1},\bar{r}_{i+1})\not\varepsilon(t_i,t)XY|\mathcal{B}_{t_i}]$$

$$= 1-\Pr[N([t_i,t)XY)= 0|\mathcal{B}_{t_i}]$$

$$= 1-\Pr[N([t_i,t)XY) = 0]$$

where the last step is due to the independent increment property of the Poisson distributed process. Also, from the Poisson statistics

$$\Pr[N([t_i,t)XY) = 0] = \exp\left[-\int_{t_i}^{t}\int_{Y}\lambda(\tau,\bar{\pi};\omega;\omega_s)d\bar{\pi}d\tau\right]$$

$$= \exp\left[-\int_{t_i}^{t}\int_{Y}\alpha_1(\tau,\bar{\pi};\omega_{s_3})\lambda_s(\tau,\bar{\pi};\omega;\omega_{s_1})d\bar{\pi}d\tau \right. \qquad (119)$$

$$\left. -\int_{t_i}^{t}\int_{Y}\alpha_0(\tau,\bar{\pi};\omega_{s_3})\lambda_n(\tau,\bar{\pi};\omega;\omega_{s_2})d\bar{\pi}d\tau\right]$$

In order to prove this condition, it is necessary to show $\Pr[N([t_i,t)XY) = 0]>0$. Equation (119) can only equal zero if the exponent is infinite. Because the rate parameter for the signal process, $\lambda_s(t,\bar{r};\omega;\omega_{s_1})$, is Gaussian shaped

$$\int_{Y} \alpha_1(t,\bar{\pi};\omega_{s_3})\lambda_s(t,\bar{\pi};\omega;\omega_{s_1})d\bar{\pi} \leq \int_{R^m} \lambda_s(t,\bar{\pi};\omega;\omega_{s_1})d\bar{\pi}$$

(120)

$$= \Lambda(t)(2\pi)^{\frac{m}{2}}|\underline{R}(t)|^{\frac{1}{2}} < \infty$$

Note that the Gaussian shaped intensity profile for the signal rate parameter yields a convenient evaluation of equation (120) and allows us to use the Snyder-Fishman filter results for the elemental estimators of the MMAE. The Gaussian shape itself, however, is not necessary, as long as the conditions of this theorem are satisfied and we use an appropriate elemental estimator. We now need the previously mentioned condition

$$\int_{t_0}^{t} \int_{Y} \alpha_0(t,\bar{\pi};\omega_{s_3})\lambda_n(t,\bar{\pi};\omega;\omega_{s_2})d\bar{\pi}\, d\tau < \infty \qquad (121)$$

Since $\alpha_0(t,\bar{r};\omega_{s_3})$ and $\alpha_1(t,\bar{r};\omega_{s_3})$ each have a maximum value of one, the entire exponent is finite for finite t. Therefore

$$\Pr[N([t_i,t)XY = 0] = \exp\left[-\int_{t_i}^{t}\int_{Y}\lambda(\tau,\bar{\pi};\omega;\omega_s)d\bar{\pi}d\tau\right] > 0 \quad (122)$$

Substitution of equation (122) into equation (118) results in

$$\Pr[(t_{i+1},\bar{r}_{i+1})\varepsilon(t_i,t)XY|\mathcal{B}_{t_i}] < 1 \quad (123)$$

w.p. 1

(c) To show that: The point process is conditionally orderly:

To simplify notation, let

$$N(\Delta t,\bar{\rho}) \triangleq N([t,t+\Delta t)\times c(\bar{r},\bar{\rho})) \tag{124}$$

With this notation, we can write the expression to be evaluated (from equation (75)) as

$$\lim_{\substack{\Delta t\to 0\\ \bar{\rho}\to 0}} \frac{Pr[N(\Delta t,\bar{\rho})\geq 2 \mid \mathcal{B}_t]}{Pr[N(\Delta t,\bar{\rho})\geq 1 \mid \mathcal{B}_t]} \tag{125}$$

$$= \lim_{\substack{\Delta t\to 0\\ \bar{\rho}\to 0}} \frac{1-Pr[N(\Delta t,\bar{\rho})=0 \mid \mathcal{B}_t]-Pr[N(\Delta t,\bar{\rho})=1 \mid \mathcal{B}_t]}{1-Pr[N(\Delta t,\bar{\rho})=0 \mid \mathcal{B}_t]}$$

$$= \lim_{\substack{\Delta t \to 0 \\ \bar{\rho} \to 0}} \frac{1-\Pr[N(\Delta t,\bar{\rho}) = 0]-\Pr[N(\Delta t,\bar{\rho}) = 1]}{1-\Pr[N(\Delta t,\bar{\rho}) = 0]} \tag{126}$$

where the last step is due to the independent increment property of the Poisson point process. Since the point process is Poisson,

$$\lim_{\substack{\Delta t \to 0 \\ \bar{\rho} \to 0}} \Pr[N(\Delta t,\bar{\rho}) = 1]$$

$$= \lim_{\substack{\Delta t \to 0 \\ \bar{\rho} \to 0}} \int_{t}^{t+\Delta t} \int_{c(\bar{r},\bar{\rho})} \lambda(\tau,\bar{\pi};\omega;\omega_s) d\bar{\pi} d\tau \tag{127}$$

$$\cdot \exp\left[- \int_{t}^{t+\Delta t} \int_{c(\bar{r},\bar{\rho})} \lambda(\tau,\bar{\pi};\omega;\omega_s) d\bar{\pi} d\tau \right] = 0$$

and

$$\lim_{\substack{\Delta t \to 0 \\ \bar{\rho} \to 0}} \Pr[N(\Delta t, \bar{\rho}) = 0]$$

$$= \lim_{\substack{\Delta t \to 0 \\ \bar{\rho} \to 0}} \exp\left[- \int_{t}^{t+\Delta t} \int_{c(\bar{r},\bar{\rho})} \lambda(\tau, \bar{\pi}; \omega; \omega_s) d\bar{\pi} d\tau \right] = 1 \tag{128}$$

therefore, when we attempt to take the limit of equation (126), the result is indeterminate

$$\frac{1-1-0}{1-1} = \frac{0}{0} \tag{129}$$

In order to evaluate equation (126), we must use L'Hospitals rule and Leibnitz's rule, with respect to ρ_1

(the first component of $\bar{\rho}$) first:

$$\lim_{\substack{\Delta t \to 0 \\ \bar{\rho} \to 0}} \frac{\Pr[N(\Delta t, \bar{\rho}) \geq 2 | \mathcal{B}_t]}{\Pr[N(\Delta t, \bar{\rho}) \geq 1 | \mathcal{B}_t]}$$

$$= \lim_{\substack{\Delta t \to 0 \\ \rho_2 \to 0 \\ \vdots \\ \rho_m \to 0}} \lim_{\rho_1 \to 0} \frac{-\frac{\partial}{\partial \rho_1} \Pr[N(\Delta t, \bar{\rho}) = 0] - \frac{\partial}{\partial \rho_1} \Pr[N(\Delta t, \bar{\rho}) = 1]}{-\frac{\partial}{\partial \rho_1} \Pr[N(\Delta t, \bar{\rho}) = 0]} \tag{130}$$

$$= \lim_{\substack{\Delta t \to 0 \\ \rho_2 \to 0 \\ \vdots \\ \rho_m \to 0}} \lim_{\rho_1 \to 0} \frac{e^{-\theta\psi} + \theta e^{-\theta\psi} - e^{-\theta\psi}}{e^{-\theta\psi}} \tag{131}$$

$$= \lim_{\substack{\Delta t \to 0 \\ \rho_2 \to 0 \\ \vdots \\ \rho_m \to 0}} \lim_{\rho_1 \to 0} \frac{\theta e^{-\theta}\psi}{e^{-\theta}\psi} \tag{132}$$

where

$$\theta \triangleq \int_t^{t+\Delta t} \int_{c(\bar{r},\bar{\rho})} \lambda(\tau,\bar{\pi};\omega;\omega_s) d\bar{\pi} d\tau \tag{133}$$

and

$$\psi = \int_{t}^{t+\Delta t} \int_{r_2}^{r_2+\rho_2} \cdots \int_{r_m}^{r_m+\rho_m} \lambda(\tau, r_1+\rho_1, \hbar_2, \hbar_3, \ldots, \hbar_m; \omega; \omega_s) d\hbar_m \ldots d\hbar_2 d\tau \tag{134}$$

From equation (133), it can be seen that

$$\lim_{\rho_1 \to 0} \theta = 0 \tag{135}$$

and the limit of equation (132) can be taken with respect to ρ_1 and we have the desired result.

$$\lim_{\substack{\Delta t \to 0 \\ \bar{\rho} \to 0}} \frac{\Pr[N(\Delta t, \bar{\rho}) \geq 2 | \mathcal{B}_t]}{\Pr[N(\Delta t, \bar{\rho}) \geq 1 | \mathcal{B}_t]} = 0 \qquad (136)$$

The process is conditionally orderly.

(d) To show that: $\Pr[N([t_0, t)XY) < \infty] = 1$ for finite t

$$\Pr[N([t_0, t)XY) < \infty] = 1 - \Pr[N([t_0, t)XY) = \infty] \qquad (137)$$

For the analytical model description under consideration

$$\Pr[N([t_0,t)XY) = n]$$

$$= \frac{\left[\int_{t_0}^{t}\int_{Y}\lambda(\tau,\bar{\pi};\omega;\omega_s)d\bar{\pi}d\tau\right]^n \exp\left[-\int_{t_0}^{t}\int_{Y}\lambda(\tau,\bar{\pi};\omega;\omega_s)d\bar{\pi}d\tau\right]}{n!} \qquad (138)$$

$$n = 0,1,2,\ldots$$

therefore, for finite t, and by equations (120) and (121)

$$\lim_{n\to\infty} \Pr[N([t_0,t)XY)=n]=\Pr[N([t_0,t)XY)=\infty]= 0 \qquad (139)$$

and

$$\Pr[N([t_0,t)XY)<\infty] = 1 - 0 = 1 \tag{140}$$

The proof is complete. □

We have shown that U is regular given $\omega_s \varepsilon \Omega_s$. The following theorem shows that, for the analytical cross product space model given above, the space-time point process V is regular. With this condition satisfied, we know that a hazard function exists for V and it can be evaluated as in equation (87).

Theorem III-8. Let the probability spaces (Ω_s, A_s, P_s) and $(\Omega, \mathcal{B}, P(\cdot;\omega_s))$ satisfy Theorem III-6. If

(a)

$$E\{\lambda_n(t,\bar{r};\omega;\omega_{s_2})\}=\lambda_n(t,\bar{r})=\int_{\Omega X\Omega_s}\lambda_n(t,\bar{r};\omega;\omega_{s_2})P'(d\omega X d\omega_{s_2})<\infty \tag{141}$$

and

(b) $\lambda(t,\bar{r};\omega;\omega_s)$ is measurable with respect to

$[t_0,t)\ X\ R^m\ X\ \Omega\ X\ \Omega_s$

then the space-time point process V is regular.

■

Proof. To prove Theorem III-8, we must show that the analytical description of the cross product spaces satisfies Theorem III-1 (equations (83), (84), and (85)).

(a) To show that: If $B\varepsilon\mathcal{B}$ and $t\varepsilon[t_0,T)$ then $P(B;\omega_s|\mathcal{B}_t)(\omega)$ is measurable with respect to the product sigma field $A_s\otimes\mathcal{B}$

Consider the function (mapping):

$$P:\Omega X\Omega_s \rightarrow [0,1]\varepsilon R^1$$

where

$$P \triangleq P(B;\omega_s|\mathcal{B}_t)(\omega)$$

is a set function which maps from the cross product space $\Omega X\Omega_s$ into a closed interval on the real line. The function P is measurable with respect to $A_s \otimes \mathcal{B}$ if the set

$$\{B X \omega_s : P(B;\omega_s|\mathcal{B}_t)(\omega) > \alpha\}$$

is measurable for all $\alpha\varepsilon[0,1]$ (Ref. 36:65). Since it is given that $B\varepsilon\mathcal{B}$ and, by definition of the probability space, $\omega_s\varepsilon\Omega_s$ (and therefore $\omega_s \varepsilon A_s$), $B X \omega_s$ is measurable with respect to $A_s \otimes \mathcal{B}$ (Ref. 32:71). Thus $P(B;\omega_s|\mathcal{B}_t)(\omega)$ is measurable with respect to $A_s \otimes \mathcal{B}$.

(b) To show that: For each point $(t,\bar{r})\varepsilon[t_0,T)XY$

$$\phi(t,\bar{r})=E\{\phi(t,\bar{r};\omega;\omega_s)\}=\int_{\Omega X\Omega_s}\phi(t,\bar{r};\omega;\omega_s)P'(d\omega\ d\omega_s)<\infty \qquad (142)$$

The hazard function $\phi(t,\bar{r};\omega;\omega_s)$ is measurable with respect to each of its arguments. By the Fubini Theorem

$$\phi(t,\bar{r}) = \int_{\Omega}\int_{\Omega_s}\phi(t,\bar{r};\omega;\omega_s)P_s(d\omega_s)P(d\omega) \qquad (143)$$

$$= \int_{\Omega}\int_{\Omega_{S_3}} \left[\alpha_1(t;\omega_{S_3}) \int_{\Omega_{S_1}} \lambda_S(t,\bar{r};\omega;\omega_{S_1}) P_{S_1}(d\omega_{S_1}) \right. \tag{144}$$

$$\left. + \alpha_0(t;\omega_{S_3}) \int_{\Omega_{S_2}} \lambda_n(t,\bar{r};\omega;\omega_{S_2}) P_{S_2}(d\omega_{S_2}) \right] P_{S_3}(d\omega_{S_3}) P(d\omega)$$

If $\lambda_S(t,\bar{r};\omega;\omega_{S_1})$ is Gaussian shaped as previously described, then

$$\int_{\Omega_{S_1}} \lambda_S(t,\bar{r};\omega;\omega_{S_1}) P_{S_1}(d\omega_{S_1}) = \lambda_S(t,\bar{r};\omega) < \infty \tag{145}$$

It is given that $\lambda_n(t,\bar{r}) < \infty$. Since $\alpha_0(t;\omega_{S_3})$ and $\alpha_1(t;\omega_{S_3})$ have a maximum value of one, we can bound $\phi(t,\bar{r})$ by

$$\phi(t,\bar{r}) \leq \int_{\Omega} \lambda_s(t,\bar{r};\omega)P(d\omega)+\lambda_n(t,\bar{r}) < \infty \qquad (146)$$

The condition is satisfied.

(c) To show that: $\phi(t,\bar{r};\omega;\omega_s)$ is measurable with respect to $[t_0,t) \times R^m \times B \times A_s$.

This condition is given in the statement of the theorem.

The proof is complete. ∎

Condition (c) for this proof was given in the statement of the theorem. This was done to keep the form of $\lambda(t,\bar{r};\omega;\omega_s)$ as general as possible to admit a large class of functional forms. In order for this theorem to be of practical value, however, some discussion on this point is in order. A rate parameter function $\lambda(t,\bar{r};\omega;\omega_s)$ is measurable on $[t_0,t) \times R^m \times B \times A_s$ if it is measurable individually with respect to each of its arguments when the other three arguments take on any allowable value.

Therefore, we can consider the arguments individually providing the value of one argument does not affect the measurability of the function with respect to another argument. From measure theory (Ref. 18:284-287, for example) we know that continuous functions, the sum of two measurable functions, the product of two measurable functions, and limits of measurable functions are all measurable functions. By using these results, we can construct a large class of useful measurable $\lambda(t,\bar{r};\omega;\omega_s)$ functions from elementary functions.

III.6 Summary

In this chapter, the basic physical model described in Chapter II is recast in terms of a doubly stochastic space-time point process defined on a cross product of two probability spaces. Several analytical tools are presented which allow inference of useful information from the observed doubly stochastic space-time point process. In particular, if the observed process is regular (Definition III-5), then its hazard function exists and the existence of the sample function density is guaranteed. Also, a representation theorem is presented which gives us the means of calculating the a posteriori probability of an event in the unobserved (Ω_s,A_s,P_s) probability space given observation of events in $(\Omega,\mathcal{B},\bar{P})$. Finally, an analytical

description of the particle beam estimation problem is given in terms of a cross product model. The resulting process V is shown to be regular.

Chapter II provided the basic form of the multiple model adaptive estimator suitable for this point process problem; however, the filter weights were difficult to calculate from a probability density function approach. The methods and results of Chapter III provide a means of calculating the individual filter weights when the analytical model meets the regularity conditions necessary for Theorems III-7 and III-8. In Chapter IV, an analytical model appropriate to the particle beam problem is presented, the individual filter weighting factor expression is evaluated, and the multiple model adaptive estimator equations are presented.

IV. The Estimator

IV.1 Introduction

In this chapter, the full scale multiple model adaptive estimator for the particle beam problem is developed. The first section develops the a posteriori filter weights. The general form of the multiple model adaptive estimator is shown in Chapter II, equation (24): an estimate is generated as the probabilistically weighted sum of outputs of individual estimators, each based on a specific assumption about uncertain parameters. The same form in notation suitable for the point process problem is given in equation (36). To derive the a posteriori individual filter weights, $Pr[h_j^{N_t} | \mathcal{B}_t]$, $j = 0,1,2, \ldots N_t-1$, the representation theorem from Chapter III is used. Section IV.3 presents the equations for the general multiple model adaptive estimator for the particle beam problem in cross product space terms. In this section, the estimator development is consolidated in one location for reference. Finally, Section IV.4 presents examples of the estimator equations under several simplifying assumptions. We begin by deriving the weighting factors of equation (36).

IV.2 The Multiple Model Weighting Factors

The purpose of this section is to evaluate the weighting factors, $\Pr[h_j^{N_t} | \mathcal{B}_t]$, for the multiple model adaptive estimator described in equation (36). In words, a weighting factor is the conditional probability that the j^{th} hypothesis $h_j^{N_t}$, out of a possible 2^{N_t} possible hypotheses, is the correct one, given the sub-sigma field of events up to time t. The representation theorem, Theorem III-6, gives us the means of calculating the probability of an event $A \varepsilon \mathcal{A}_s$ given $\mathcal{B}_t$. If we let the event A_j be the event associated with the hypothesis sequence $h_j^{N_t}$, then from equation (95) we have

$$\Pr[h_j^{N_t} | \mathcal{B}_t] = \frac{\int_{A_j} L_t(\omega;\omega_s) P_s(d\omega_s)}{\int_{\Omega_s} L_t(\omega;\omega_s) P_s(d\omega_s)} \tag{147}$$

where $L_t(\omega;\omega_s)$ is defined by equation (93).

Since we are using a cross product of three probability spaces to model the observed event, $L_t(\omega;\omega_s)$ is actually dependent on the events in all three probability spaces: $L_t(\omega;\omega_{s_1};\omega_{s_2};\omega_{s_3})$. The hazard functions are assumed to be of the form

$$\phi(t,\bar{r};\omega;\omega_s) = \alpha_1(t,\bar{r};\omega_{s_3})\phi_s(t,\bar{r};\omega;\omega_{s_1})$$

(148)

$$+\alpha_0(t,\bar{r};\omega_{s_3})\phi_n(t,\bar{r};\omega;\omega_{s_2})$$

Note that this form is very similar to that shown in equation (102); however, for the development of the weighting factors, we do not need to assume that the observed process is Poisson nor do we need to assume the Gaussian shaped intensity profile for

$$\phi_s(t,\bar{r};\omega;\omega_{s_1})$$

We only require that the process V be regular (that is, it satisfies Theorem III-8).

The events A_j, j=1, ..., 2^{N_t} are, by construction, mutually exclusive and exhaustive, thus

$$\bigcup_{j=1}^{2^{N_t}} A_j = \Omega_{s_3} \tag{149}$$

Because of this, we can expand $\bar{L}_t(\omega)$ as

$$\bar{L}_t(\omega) \triangleq \int_{\Omega_s} L_t(\omega;\omega_s) P_s(d\omega_s) \tag{150}$$

$$= \int_{\Omega_{s_1}} \int_{\Omega_{s_2}} \int_{\Omega_{s_3}} L_t(\omega;\omega_{s_1};\omega_{s_2};\omega_{s_3}) P_{s_3}(d\omega_{s_3}) P_{s_2}(d\omega_{s_2}) P_{s_1}(d\omega_{s_1})$$

$$= \int_{\Omega_{s_1}} \int_{\Omega_{s_2}} \sum_{j=1}^{2^{N_t}} \int_{A_j} L_t(\omega;\omega_{s_1};\omega_{s_2};\omega_{s_3}) \tag{152}$$

$$\bullet P_{s_3}(d\omega_{s_3}) P_{s_2}(d\omega_{s_2}) P_{s_1}(d\omega_{s_1})$$

$$= \sum_{j=1}^{2^{N_t}} \int_{A_j} L_t(\omega;\omega_s) P_s(d\omega_s) \tag{153}$$

In equation (153), the A_j notation is used to indicate integration over $\Omega_{s_1} X \Omega_{s_2} X A_j$, therefore

$$\sum_{j=1}^{2^{N_t}} \Pr[h_j^{N_t} | \mathcal{B}_t] = \frac{\sum_{j=1}^{2^{N_t}} \int_{A_j} L_t(\omega;\omega_s) P_s(d\omega_s)}{\bar{L}_t(\omega)} \tag{154}$$

$$= \frac{\bar{L}_t(\omega)}{\bar{L}_t(\omega)} = 1 \tag{155}$$

as expected.

By equation (94), the denominator of equation (147) is

$$\int_{\Omega_s} L_t(\omega;\omega_s) P_s(d\omega_s) = \bar{L}_t(\omega) \tag{156}$$

$$= \prod_{i=1}^{N_t} \hat{\phi}(t_i, \bar{r}_i;\omega) \exp\left[-\int_{t_0}^{t} \iint_{Y} \hat{\phi}(\tau, \bar{\pi};\omega) d\bar{\pi} d\tau \right]$$

$$= \prod_{i=1}^{N_t} [\hat{\phi}_s(t_i,\bar{r}_i;\omega)+\hat{\phi}_n(t_i,\bar{r}_i;\omega)] \qquad (157)$$

$$\cdot \exp\left[-\int_{t_0}^{t}\int_{Y}\hat{\phi}(\tau,\bar{\pi};\omega)d\bar{\pi}d\tau\right]$$

where the last step is due to the fact that there is no conditioning on A_j therefore

$$\alpha_0(t,\bar{r};\omega_{s_3}) = \alpha_1(t,\bar{r};\omega_{s_3}) = 1 \qquad (158)$$

We must next evaluate the numerator of equation (147) to obtain the weighting factors. The numerator can be expanded as

$$\int_{A_j} L_t(\omega;\omega_s)P_s(d\omega_s)$$

$$= \int_{\Omega_{s_1}} \int_{\Omega_{s_2}} \int_{\Omega_{s_3}} I_j(\omega_{s_3})L_t(\omega;\omega_{s_1};\omega_{s_2};\omega_{s_3}) \qquad (159)$$

$$\cdot P_{s_3}(d\omega_{s_3})P_{s_2}(d\omega_{s_2})P_{s_1}(d\omega_{s_1})$$

where $I_j(\omega_{s_3})$ is an indicator function defined by

$$I_j(\omega_{s_3}) = \begin{cases} 1 & \omega_{s_3} \in A_j \\ 0 & \text{otherwise} \end{cases} \qquad (160)$$

We can rewrite the integrand as

$$I_j(\omega_{s_3})L_t(\omega;\omega_s) = L_t(\omega;\omega_s)\Big|_{A_j}$$

$$= \prod_{S_j} \phi_s(t_k,\bar{r}_k;\omega;\omega_{s_1}) \prod_{N_j} \phi_n(t_\ell,\bar{r}_\ell;\omega;\omega_{s_2})$$

$$\cdot \exp\left[-\int_{t_0}^{t}\int_{Y}\phi(\tau,\bar{\pi};\omega;\omega_s)d\bar{\pi}d\tau\right] \quad (161)$$

where the vertical bar is read as "restricted to A_j ", the set of signal indices is defined as

$$S_j = \{k:h_j^{N_t}(k)=1\}=\{k_1,k_2, \ldots, k_q\} \quad (162)$$

and the set of noise indices is defined as

$$N_j = \{\ell : h_j^{N_t}(\ell)=0\}=\{\ell_1,\ell_2,\ldots,\ell p\} \tag{163}$$

and $q+p=N_t$. Note that the value of the exponential term in equation (161) is not a function of ω_{s_3} . This is because the α terms in $\phi(t,\bar{r};\omega;\omega_s)$ only have values of zero at points (sets of measure zero). We can therefore substitute equation (161) into equation (159) to obtain

$$\int_{A_j} L_t(\omega;\omega_s)P_s(d\omega_s)$$

$$= \int_{\Omega_{s_1}}\int_{\Omega_{s_2}} \prod_{S_j} \phi_s(t_k,\bar{r}_k;\omega;\omega_{s_1}) \prod_{N_j} \phi_n(t_\ell,\bar{r}_\ell;\omega;\omega_{s_2})$$

$$\cdot \exp\left[-\int_{t_0}^{t}\int_{Y} \phi(\tau,\bar{\pi};\omega;\omega_{s_1};\omega_{s_2})d\bar{\pi}d\tau\right] P_{s_2}(d\omega_{s_2})P_{s_1}(d\omega_{s_1}) \tag{164}$$

In order to evaluate equation (164), consider the process

$$V_j : \Omega_{S_1} X \Omega_{S_2} X \Omega_{S_3} X [t_0, t) \rightarrow [t_0, t) X R^m$$

where there is only one possible hypothesis sequence of signal/noise observed events; that is

$$P_{S_3}(A_j) = 1 \tag{165}$$

The sequence $h_j^{N_t}$ is guaranteed to occur. The process V_j is regular because it is a special case of the process V_i which is regular. Since V_j is regular, by Theorem III-3 a hazard function exists which is an expected value:

$$\bar{\phi}_j(t, \bar{r}; \omega) = \phi_j(t, \bar{r}; \omega) = E_S\{\phi(t, \bar{r}; \omega; \omega_S) | A_0 \otimes \delta_t\} \tag{166}$$

in which the j subscripts indicate association with the

special case process V_j . By Theorem III-5, equation (94),

$$\bar{L}_{t_j}(\omega) = \prod_{i=1}^{N_t} \hat{\phi}_j(t_i, \bar{r}_i; \omega) \exp\left[-\int_{t_0}^{t} \int_{Y} \hat{\phi}_j(\tau, \bar{\pi}; \omega) d\bar{\pi} d\tau \right] \tag{167}$$

$$= \int_{\Omega_s} L_{t_j}(\omega; \omega_s) P_s(d\omega_s) \tag{168}$$

where $L_{t_j}(\omega;\omega_s)$ is the sample function density for V_j as defined by equation (93):

$L_{t_j}(\omega;\omega_s) =$

$$\prod_{i=1}^{N_t} \phi_j(t_i, \bar{r}_i; \omega; \omega_s) \exp\left[-\int_{t_0}^{t} \int_{Y} \phi_j(\tau, \bar{\pi}; \omega; \omega_s) d\bar{\pi} d\tau \right] \tag{169}$$

The sure event A_j determines the value of the coefficients

$\alpha_0(t,\bar{r};\omega_{s_3})$ and $\alpha_1(t,\bar{r};\omega_{s_3})$; therefore

$$L_{t_j}(\omega;\omega_s) = \prod_{S_j} \phi_s(t_k,\bar{r}_k;\omega;\omega_{s_1}) \prod_{N_j} \phi_n(t_\ell,\bar{r}_\ell;\omega;\omega_{s_2})$$

$$\exp\left[-\int_{t_0}^{t}\int_{Y} \phi_j(\tau,\bar{\pi};\omega;\omega_s)d\bar{\pi}d\tau \right] \qquad (170)$$

We can substitute equation (170) into equation (168) to obtain

$$\bar{L}_{t_j}(\omega) = \int_{\Omega_s} L_{t_j}(\omega;\omega_s)P_s(d\omega_s) \qquad (171)$$

$$= \int_{\Omega_{s_1}} \int_{\Omega_{s_2}} \prod_{S_j} \phi_s(t_k, \bar{r}_k; \omega; \omega_{s_1}) \prod_{N_j} \phi_n(t_\ell, \bar{r}_\ell; \omega; \omega_{s_2})$$

$$\cdot \exp\left[- \int_{t_0}^{t} \int_{Y} \phi(\tau, \bar{\pi}; \omega; \omega_{s_1}; \omega_{s_2}) d\bar{\pi} d\tau \right]$$

$$\cdot P_{s_2}(d\omega_{s_2}) P_{s_1}(d\omega_{s_1}) \qquad (172)$$

where we have used the property that the integral term in the exponential of equation (172) is independent of ω_{s_3} because the α coefficients of $\phi(\cdot)$ only have a value of zero at points. Equation (172) is exactly the expression we need to evaluate in equation (164). By equation (168), we know that

$$\bar{L}_{t_j}(\omega)$$

$$= \prod_{S_j} \hat{\phi}_s(t_k, \bar{r}_k, \omega) \prod_{N_j} \hat{\phi}_n(t_\ell, \bar{r}_\ell; \omega) \exp\left[-\int_{t_0}^{t} \int_{Y} \hat{\phi}(\tau, \bar{\pi}; \omega) d\bar{\pi} d\tau \right] \tag{173}$$

We can substitute equation (173) into equation (164) to arrive at the numerator:

$$\int_{A_j} L_t(\omega; \omega_s) P_s(d\omega_s)$$

$$= \prod_{S_j} \hat{\phi}_s(t_k, \bar{r}_k; \omega) \prod_{N_j} \hat{\phi}_n(t_\ell, \bar{r}_\ell; \omega) \exp\left[-\int_{t_0}^{t} \int_{Y} \hat{\phi}(\tau, \bar{\pi}; \omega) d\bar{\pi} d\tau \right] \tag{174}$$

Substitution of equations (174) and (157) into equation (147) results in the desired weighting factor

$$\Pr[h_j^{N_t} | \mathcal{B}_t] = \frac{\prod_{S_j} \hat{\phi}_s(t_k, \bar{r}_k; \omega) \prod_{N_j} \hat{\phi}_n(t_\ell, \bar{r}_\ell; \omega)}{\prod_{i=1}^{N_t} \left[\hat{\phi}_s(t_i, \bar{r}_i; \omega) + \hat{\phi}_n(t_i, \bar{r}_i; \omega) \right]} \tag{175}$$

From this expression for the weighting factors, it can be seen that

$$\sum_{j=2}^{2^{N_t}} \Pr[h_j^{N_t} | \mathcal{B}_t] = 1 \tag{176}$$

as discussed in the text preceding equation (155). Furthermore, the weighting factors are recursive:

$$\Pr[h_j^{N_t} | \mathcal{B}_t] = \frac{h_j^{N_t}(N_t)\hat{\phi}_s(t_{N_t}, \bar{r}_{N_t};\omega) + \{1-h_j^{N_t}(N_t)\}\hat{\phi}_n(t_{N_t}, \bar{r}_{N_t};\omega)}{\hat{\phi}_s(t_{N_t}, \bar{r}_{N_t};\omega) + \hat{\phi}_n(t_{N_t}, \bar{r}_{N_t};\omega)}$$

$$\cdot \Pr\left[h_{j'}^{N_t-1} | \mathcal{B}_{t_{(N_t-1)}}\right]$$

$$\text{for } t_{N_t} \leq t \leq t_{N_t+1} \qquad (177)$$

where the j' notation denotes the "old" sequence which is concatenated with the new sequence value, $h_j^{N_t}(N_t)$, to obtain the "new" sequence $h_j^{N_t}$. This is the same notation discussed in more detail in the text following equation (51).

We now have the weighting coefficients $\Pr[h_j^{N_t} | \mathcal{B}_t]$ for use in the multiple model adaptive estimator equation

(36). The evaluation of the coefficients is general because we did not need to specify the exact form of the hazard function. It was only necessary that the observed process satisfy Theorem III-8. In the next section, all the equations for the multiple model adaptive estimator for the particle beam problem are consolidated and presented.

IV.3 Estimator Equations

In this section, the results of the previous section and chapters are consolidated and the multiple model adaptive estimator equations for the particle beam problem are presented. Only the results are listed here; references to earlier chapters and sections are provided for detailed descriptions and derivations.

IV.3.1 Assumptions. A conditionally Poisson space-time point process is observed on $[t_0,t)\times R^m$. The observed process consists of point process events from a conditionally Poisson signal process with rate parameter $\lambda_s(t,\bar{r};\omega;\omega_{s_1})$ and a conditionally Poisson noise process with rate parameter $\lambda_n(t,\bar{r};\omega;\omega_{s_2})$. The signal and noise processes are assumed independent, therefore the rate parameter for the observed process is

$$\lambda(t,\bar{r};\omega;\omega_s)=\lambda_s(t,\bar{r};\omega;\omega_{s_1})+\lambda_n(t,\bar{r};\omega;\omega_{s_2}) \tag{178}$$

The signal rate parameter is assumed to have a Gaussian shape as defined by equation (103) and the noise rate parameter is defined by equation (105). We further assume that

$$\int_{t_0}^{t}\int_{R^m} \lambda_n(\tau,\bar{\pi};\omega;\omega_{s_2})d\bar{\pi}d\tau < \infty \tag{179}$$

$$E\{\lambda_n(t,\bar{r};\omega;\omega_{s_2})\} < \infty$$

and that

$$d\underset{\sim}{\bar{x}}(t) = \underline{F}(t)\underset{\sim}{\bar{x}}(t)dt + \underline{G}(t)d\underset{\sim}{\bar{u}}(t) \tag{180}$$

as in Chapter II.

IV.3.2 The Multiple Model Adaptive Estimator. The multiple model adaptive estimate of $\bar{x}(t;\omega_{s_1})$ is given by

$$\hat{\bar{x}}(t;\omega_{s_1}) = \sum_{j=0}^{2^{N_t}-1} \hat{\bar{x}}_j(t) Pr[h_j^{N_t} | \mathcal{B}_t] \tag{181}$$

where N_t is the number of observed events in $[t_0,t)XR^m$. The development of the MMAE form is presented in Chapter II. The individual hypothesis estimates, $\hat{\bar{x}}_j(t)$, are defined as

$$\hat{\bar{x}}_j(t) \triangleq E\{\bar{x}(t;\omega_{s_1}) | \mathcal{B}_t, h_j^{N_t}\} \tag{182}$$

where $\mathcal{B}_t$ is the sub-sigma field of events up to time t and $h_j^{N_t}$ is a hypothesis sequence which defines the j^{th} entry out of a list of 2^{N_t} possible hypothesis sequences at time t. The hypothesis sequence notation is defined in

detail following equation (28).

The process $\underset{\sim}{\bar{x}}(t)$ is modeled as the output of a linear system driven by unit-strength white Gaussian noise. We observe a conditionally Poisson point process which has a Gaussian shaped rate parameter. As a result of this model, we can use the Snyder-Fishman filter to calculate the individual estimates.

The general form of the Snyder-Fishman filter equations used to calculate the $\hat{\bar{x}}_j$ terms is given by equations (10-13). Since the hypothesis sequence $h_j^{N_t}$ is given (in equation (182)), the estimate for $\hat{\bar{x}}_j$ is calculated by considering only those observed events which are caused (according to $h_j^{N_t}$) by the underlying signal process. Thus, for each of the 2^{N_t} models, we calculate

$$d\hat{\bar{x}}_j(t) = \underline{F}(t)\hat{\bar{x}}_j(t)dt + \int_{R^m} \underline{K}_j(t)\left[\bar{r} - \underline{H}(t)\hat{\bar{x}}_j(t)\right] N(dt \times d\bar{r}) \quad (183)$$

$$\left\{ + \bar{c}'(t) \text{ if control is included as in equation (104)} \right\}$$

and

$$d\hat{\underline{\Sigma}}_j(t) = \underline{F}(t)\hat{\underline{\Sigma}}_j(t)dt+\hat{\underline{\Sigma}}_j(t)\underline{F}^T(t)dt+\underline{G}(t)\underline{G}(t)^T(t)dt$$

$$-\int_{R^m}\underline{K}_j(t)\underline{H}(t)\hat{\underline{\Sigma}}_j(t)N(dt\times d\bar{\kappa}) \tag{184}$$

where the counting integrals in equations (183) and (184) only have nonzero value at the measurements $(t,\bar{r})$ which are specified by $h_j^{N_t}$ to have been caused by the signal process.

Equations (183) and (184) can also be written in propagate-update form, as is commonly done for Kalman filter realizations. Let t^- denote a time immediately prior to an observed point process event and let t^+ denote a time just after an observed point process event and its incorporation into the estimate. The estimate of $\bar{x}(t)$ is propagated by

$$d\hat{\bar{x}}_j(t) = \underline{F}(t)\hat{\bar{x}}_j(t)dt \quad \{+\ \bar{c}' \text{ if used}\}$$

$$\text{for: } t_{i-1}^{+} \leq t < t_i^{-} \tag{185}$$

$$i = 1,2,\ldots, N_t$$

Updates, due to observed point process events, are accomplished by

$$\hat{\bar{x}}_j(t_i^{+}) = \hat{\bar{x}}_j(t_i^{-}) + \underline{K}_j(t_i^{-})(\bar{r}_i - \underline{H}(t_i^{-})\hat{\bar{x}}_j(t_i^{-})) \tag{186}$$

$$i = 1,2, \ldots, N_t$$

The propagate form for the individual covariances is

$$d\hat{\underline{\Sigma}}_j(t) = \underline{F}(t)\hat{\underline{\Sigma}}_j(t)dt+\hat{\underline{\Sigma}}_j\underline{F}^T(t)dt+\underline{G}(t)\underline{G}^T(t)dt \tag{187}$$

$$t_{i-1}^+ \leq t \leq t_i^-$$

$$i = 1,2, \ldots, N_t$$

and the update expression is

$$\hat{\underline{\Sigma}}_j(t_i^+) = \hat{\underline{\Sigma}}_j(t_i^-)-\underline{K}_j(t_i^-)\underline{H}(t_i^-)\hat{\underline{\Sigma}}_j(t_i^-) \tag{188}$$

$$i=1,2, \ldots, N_t$$

where

$$\underline{K}_j(t) = \hat{\underline{\Sigma}}_j(t)\underline{H}^T(t)[\underline{H}(t)\hat{\underline{\Sigma}}_j(t)\underline{H}^T(t)+\underline{R}(t)]^{-1} \tag{189}$$

Note that we could use another (perhaps nonlinear) model for $\bar{x}(t)$ and an arbitrary rate parameter for the conditionally Poisson signal process. The general form of the MMA estimator (equation (181)) is unchanged; however, the individual estimates of $\bar{x}(t)$ would have to be rederived for the new model. The Gauss-Markov nature of $\bar{x}(t)$ and the Gaussian shaped rate parameter allow us to use the convenient Snyder-Fishman filter.

The weighting factors in equation (181) are generated as

$$\Pr[h_j^{N_t}|\beta_t] = \frac{\prod^{S_j}\hat{\lambda}_s(t_k,\bar{r}_k;\omega) \cdot \prod^{N_j}\hat{\lambda}_n(t_\ell,\bar{r}_\ell;\omega)}{\prod_{i=1}^{N_t}\left[\hat{\lambda}_s(t_i,\bar{r}_i;\omega) + \hat{\lambda}_n(t_i,\bar{r}_i;\omega)\right]} \tag{190}$$

where

$$\hat{\lambda}_s(t,\bar{r};\omega) \triangleq E_{s_1}\left\{\lambda_s(t,\bar{r};\omega;\omega_{s_1})|\mathcal{B}_t\right\} \tag{191}$$

and

$$\hat{\lambda}_n(t,\bar{r};\omega) \triangleq E_{s_2}\left\{\lambda_n(t,\bar{r};\omega;\omega_{s_2})|\mathcal{B}_t\right\} \tag{192}$$

These weighting factors are developed in Section IV.2.

From equation (27), the covariance of the estimator is

$$\hat{\underline{\Sigma}}(t) = \sum_{j=0}^{2^{N_t}-1} \Pr[h_j^{N_t}|\mathcal{B}_t] \cdot\left\{\hat{\underline{\Sigma}}_j(t)+[\hat{\bar{x}}_j(t)-\hat{\bar{x}}(t)][\hat{\bar{x}}_j(t)-\hat{\bar{x}}(t)]^T\right\} \tag{193}$$

where the individual covariances are given by the solution to equation (184) for $j = 0,1, \ldots ,2^{N_t}-1$.

The equations in this section specify the multiple model adaptive estimator for the (very general) form of $\lambda_n(t,\bar{r};\omega;\omega_{s_2})$ given. The following section describes several examples of the estimator's equations for various assumed forms of λ_s and λ_n .

IV.4 Examples

The following examples provide insight into the structure of the estimator and a sample description of the model to match a simple particle beam tracking application.

IV.4.1 Example a: Structure Insight. In this example, we see that the weighting factors defined by equation (190) reduce to equation (67) when the same assumptions are made.

For the development of equation (67), it is assumed that both the signal and noise rate parameters are nonrandom:

$$\lambda_s(t,\bar{r},\bar{x}(t;\omega_{s_1});\omega) = \lambda_s(t,\bar{r},\bar{x}(t)) \qquad (194)$$

where $\bar{x}(t)$ is known, and

$$\lambda_n(t,\bar{r};\omega;\omega_{s_2}) = \lambda_n(t,\bar{r}) \tag{195}$$

We can substitute these values into equations (191) and (192) to obtain

$$\hat{\lambda}_s(t,\bar{r};\omega) = \lambda_s(t,\bar{r}) \tag{196}$$

and

$$\hat{\lambda}_n(t,\bar{r};\omega) = \lambda_n(t,\bar{r}) \tag{197}$$

We can substitute equations (196) and (197) into equation (190) to obtain

$$\Pr[h_j^{N_t} | \mathcal{B}_t] = \frac{\prod_{S_j} \lambda_s(t_k, \bar{r}_{k'}, \bar{x}(t_k)) \prod_{N_j} \lambda_n(t_\ell, \bar{r}_\ell)}{\prod_{i=1}^{N_t} \left\{ \lambda_s(t_i, \bar{r}_i, \bar{x}(t_i)) + \lambda_n(t_i, \bar{r}_i) \right\}} \tag{198}$$

Equation (198) is the same as equation (67). Thus, for the trivial case of known $\lambda_s(t,\bar{r},\bar{x}(t))$ and $\lambda_n(t,\bar{r})$, the full scale estimator reduces to the same form as in our initial multiple model adaptive estimator development.

IV.4.2 Example b: Tracking Model. In this example, we present parameters for the model which are suitable for a simple tracking application. The problem is to estimate the position of a circularly symmetric Gaussian intensity profile on a two dimensional detector array based on the observation of space-time point process events. There is no feedback control in this example.

Let the point process be observed on $[t_0,t)XR^2$ where

$n = 2$	dimension of $\bar{x}(t)$
$m = 2$	dimension of observed spatial vectors.
$\lambda_s(t,\bar{r};\omega;\omega_{s_1})=\lambda_s(t,\bar{r};\omega_{s_1})$	
$\underline{H}(t) = \underline{H} = \underline{I}$	constant
	$\underline{I}$ = identity matrix
$\Lambda(t) = \Lambda$	constant
$\underline{R}(t) = \underline{R} = \gamma^2\underline{I}$	
γ	scalar constant
$\underline{F}(t) = -\underline{I}$	
$G(t) = g\underline{I}$	
g	scalar constant
$\underline{\tilde{u}}(t)$ = two dimensional Weiner process of unit diffusion	
$\lambda_n(t,\bar{r};\omega;\omega_{s_2})=\lambda_n(\bar{r})=\begin{cases}\lambda_n & \bar{r}\epsilon Y\subset R^2\\ 0 & \text{elsewhere}\end{cases}$	(199)
	where Y represents a two dimensional photo-detector array as in Figure 7.
$\hat{\bar{x}}(0) = \bar{0}$	initial conditions
$\underline{\Sigma}(0) = \underline{\Sigma}_0$	

The vector $\bar{x}(t;\omega_{s1})$ is a two dimensional vector output of the linear system

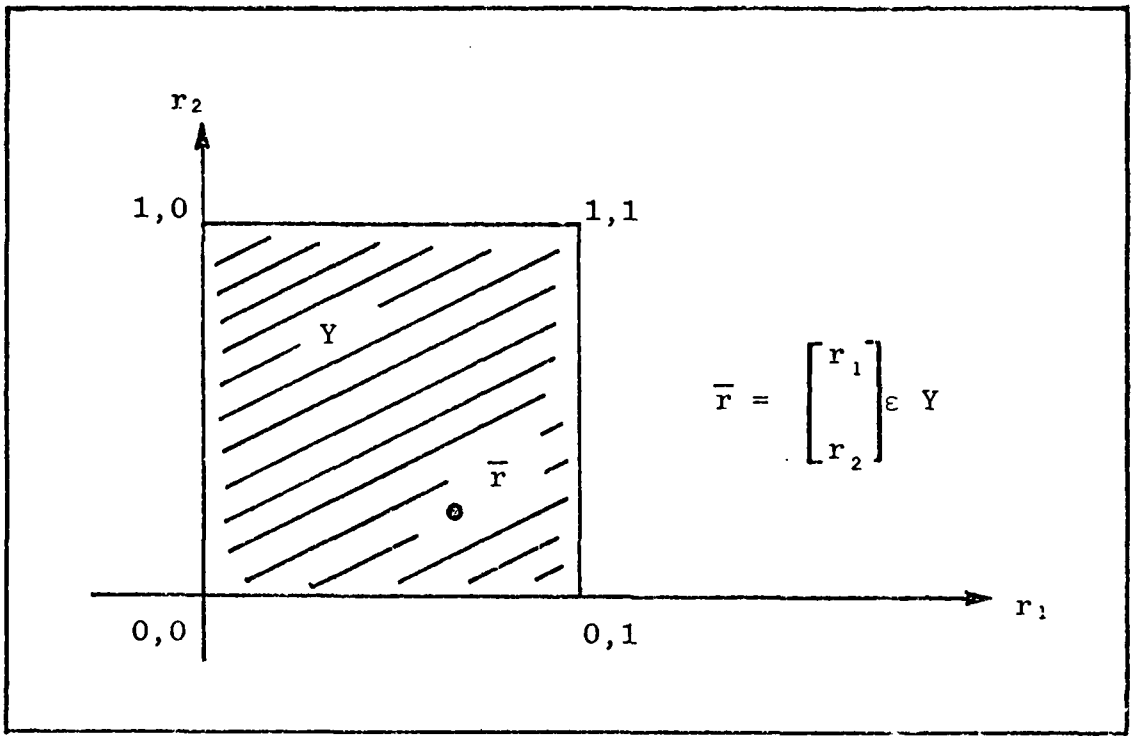


Figure 7. The Detector Model

$$d\bar{x}(t) = -\begin{bmatrix} x_1(t)dt \\ \\ x_2(t)dt \end{bmatrix} + g \begin{bmatrix} dU_1(t) \\ \\ dU_2(t) \end{bmatrix} \qquad (200)$$

The signal rate parameter is

$$\lambda_s(t,\bar{r};\omega_{s_1}) = \Lambda\exp\{-\tfrac{1}{2}[\bar{r}-\underset{\sim}{\bar{x}}(t)]^T \frac{1}{\gamma^2}\underline{I}[\bar{r}-\underset{\sim}{\bar{x}}(t)]\} \qquad (201)$$

Thus, the point of maximum intensity on the detector array is $\underset{\sim}{\bar{x}}(t)$ and our goal is to estimate $\underset{\sim}{\bar{x}}(t)$ given a sequence of point process observations of the form $(t,\bar{r})$ which are corrupted by point process noise.

The noise rate parameter is selected to model noise from dark current electron emissions in a physical photodetector. Because of the finite nature of Y, λ_n is integrable on R^2 (as required in Chapter III) and there will be no dark current induced noise events outside of the limits of Y .

Although this description of $\lambda_n(t,\bar{r})$ models the finite size of a physical detector, notice that $\lambda_s(t,\bar{r};\omega_{s_1})$ is defined for all $\bar{r}\epsilon R^m$ throughout the development of the estimator. If we simulate this tracking problem (on a computer for example) we must consider this fact. One method for dealing with this is to make the size of Y very large with respect to the "spot size" of the signal controlled by $\underline{R}$. When this is done, we can insure that

there will be an acceptably small probability of observing a signal event where there is no detector array. Note that this modeling concern does not change the estimator's equation; it only affects how accurately simulation results can be applied to a real world system and how closely any calculated error statistics match the real world system performance. A second method is not to allow any "signal" point process observed events outside of the subspace Y. This would involve a modification of the individual estimators to include the edge effects from the disallowed regions.

The expression for the estimator is given by equation (181) and the individual hypothesis estimates, $\hat{\underline{x}}_j$, are given in differential form for $j=0,1,2, \ldots ,N_t-1$ by

$$d\hat{\underline{x}}_j(t) = -\hat{\underline{x}}_j(t)dt + \int_{R^2} \underline{K}_j(t)[\underline{r}-\hat{\underline{x}}_j]N(dt\times d\underline{r}) \qquad (202)$$

$$d\hat{\underline{\Sigma}}_j(t) = -2\hat{\underline{\Sigma}}_j(t)dt + g^2\underline{I}dt - \int_{R^2} \underline{K}_j(t)\hat{\underline{\Sigma}}_j(t)N(dt \times d\bar{\pi}) \qquad (203)$$

$$\underline{K}_j(t) = \hat{\underline{\Sigma}}_j(t)[\hat{\underline{\Sigma}}_j(t) + \gamma^2\underline{I}]^{-1} \qquad (204)$$

The weighting factors are given by equation (190) where

$$\hat{\lambda}_s(t_i, \bar{r}_i) = \int_{R^2} \lambda_s(t_i, \bar{r}_i, \bar{\xi}) f_{\underline{\tilde{x}}(t_i)|\mathcal{B}_t}(\bar{\xi}|\mathcal{B}_t)d\bar{\xi} \qquad (205)$$

When the first point process event is observed at $(t_1, \bar{r}_1)$, we have

$$\Pr[h_0^1|\mathcal{B}_t] = \frac{\lambda_n(t_1,\bar{r}_1)}{\hat{\lambda}_s(t_1,\bar{r}_1)+\lambda_n(t_1,\bar{r}_1)} \tag{206}$$

$$\Pr[h_1^1|\mathcal{B}_t] = \frac{\hat{\lambda}_s(t_1,\bar{r}_1)}{\hat{\lambda}_s(t_1,\bar{r}_1)+\lambda_n(t_1,\bar{r}_1)} \tag{207}$$

Note that

$$\Pr[h_0^1|\mathcal{B}_t] + \Pr[h_1^1|\mathcal{B}_t] = 1 \tag{208}$$

This is expected since there are only two possible hypothesis sequences at time t_1 and one or the other must be correct.

At this point, it is useful to introduce a superscript for $\hat{\underline{x}}$ which is similar to that used for the hypothesis sequence notation. We let $\hat{\underline{x}}_j^{N_t}(t_i,\cdot)$ be the individual estimate of $\underline{\bar{x}}$ at time t_i, assuming that $h_j^{N_t}$ is the correct sequence. The superscript specifies that this value for $\hat{\underline{x}}$ is evaluated after incorporation of the observation numbered N_t. Note that j can take on values in the set $0,1,2, \ldots ,2^{N_t}-1$ and that $i \geq N_t$. The same superscript can be assigned to the individual covariance values: $\hat{\underline{\Sigma}}_j^{N_t}(t)$. This additional notation is necessary because the growing number of hypothesis sequences causes the subscript indices to change as each new data point is observed. This can be seen in the rest of this example.

With this notation, the individual estimates for $\hat{\underline{x}}_0^1$ and $\hat{\underline{x}}_1^1$ (in differential form) are given by

$$d\hat{\underline{x}}_0^1(t) = -\hat{\underline{x}}_0^1(t)dt \qquad (209)$$

$$t_0 \leq t \leq t_1^-$$

$$d\hat{\underline{x}}_1^1(t) = -\hat{\underline{x}}_1^1 dt \tag{210}$$

$$t_0 \leq t \leq t_1^-$$

$$\hat{\underline{x}}_1^1(t_1^+) = \hat{\underline{x}}_1^1(t_1^-)+\underline{K}_1(t_1)[\overline{r}_1-\hat{\underline{x}}_1^{1\prime}(t_1^-)] \tag{211}$$

$$d\hat{\underline{\Sigma}}_1^1(t) = -2\hat{\underline{\Sigma}}_1^1(t)dt+g^2\underline{I}dt \tag{212}$$

$$t_0 \leq t \leq t_1^-$$

$$\hat{\underline{\Sigma}}_1^1(t_1^+) = \hat{\underline{\Sigma}}_1^1(t_1^-)-\underline{K}_1(t_1^-)\hat{\underline{\Sigma}}_1^1(t_1^-) \tag{213}$$

$$\underline{K}_1(t) = \hat{\underline{\Sigma}}_1^{1}(t)[\hat{\underline{\Sigma}}_1^{1}(t)+\gamma^2\underline{I}]^{-1} \tag{214}$$

Note that we did not need to calculate $d\hat{\underline{\Sigma}}_0^{1}(t)$ explicitly for the estimate of $\hat{\hat{x}}(\cdot)$, although it is necessary for calculation of the overall covariance of the estimator.

When the second point process event is observed at $(t_2,\bar{r}_2)$, the weighting factors are

$$\Pr[h_0^2|\mathcal{B}_t] = \frac{\lambda_n(t_1,\bar{r}_1)\lambda_n(t_2,\bar{r}_2)}{\Delta} \tag{215}$$

$$\Pr[h_1^2 | \mathcal{B}_t] = \frac{\lambda_n(t_1,\bar{r}_1)\ \hat{\lambda}_s(t_2,\bar{r}_2)}{\Delta} \tag{216}$$

$$\Pr[h_2^2 | \mathcal{B}_t] = \frac{\hat{\lambda}_s(t_1,\bar{r}_1)\ \lambda_n(t_2,\bar{r}_2)}{\Delta} \tag{217}$$

$$\Pr[h_3^2 | \mathcal{B}_t] = \frac{\hat{\lambda}_s(t_1,\bar{r}_1)\hat{\lambda}_s(t_2,\bar{r}_2)}{\Delta} \tag{218}$$

where

$$\Delta \triangleq \prod_{i=1}^{2} [\hat{\lambda}_s(t_i,\bar{r}_i)+\lambda_n(t_i,\bar{r}_i)] \tag{219}$$

Note, again, that the weighting factors sum to one.

This example displays how the MMAE algorithm progresses as each new space-time point process event is observed. In general, to advance from time t_i^+ to time t_{i+1}^+ the estimator must:

(a) For $j = 2^i - 1$ down to

(1) Propagate $\hat{\hat{x}}_j^{N_t}(t_i^+)$ to $\hat{\hat{x}}_{2j}^{N_t}(t_{i+1}^-)$

(2) Propagate $\hat{\underline{\Sigma}}_j^{N_t}(t_i^+)$ to $\hat{\underline{\Sigma}}_{2j}^{N_t}(t_{i+1}^-)$

(b) For $j = 2^{i+1} - 1$ down to 0:

If j is even, then (it was a noise event)

(1) Let $\hat{\hat{x}}_j^{N_t}(t_{i+1}^+) = \hat{\hat{x}}_j^{N_t}(t_{i+1}^-)$

(2) Let $\hat{\underline{\Sigma}}_j^{N_t}(t_{i+1}^+) = \hat{\underline{\Sigma}}_j^{N_t}(t_{i+1}^-)$

Else (it was a signal event)

(3) Let $\hat{\hat{x}}_j^{N_t}(t_{i+1}^-) = \hat{\hat{x}}_{j-1}^{N_t}(t_{i+1}^-)$

(4) Update $\hat{\hat{x}}_j^{N_t}(t_{i+1}^-)$ to $\hat{\hat{x}}_j^{N_t}(t_{i+1}^+)$

(5) Let $\hat{\underline{\Sigma}}_j^{N_t}(t_{i+1}^-) = \hat{\underline{\Sigma}}_{j-1}^{N_t}(t_{i+1}^-)$

(6) Update $\hat{\underline{\Sigma}}_j^{Nt}(t_{i+1}^-)$ to $\hat{\underline{\Sigma}}_j^{Nt}(t_{i+1}^+)$

(Note the changing subscripts and the doubling of the number of active filters)

(c) For $j = 0$ to $2^{i+1} - 1$:

Calculate $Pr[h_j^{i+1} | \beta_t]$

(d) Calculate $\hat{\bar{x}}(t_{i+1})$

Figures 8 and 9 depict the storage requirements and data movements of the x vector for one iteration of the above algorithm. Figure 8 shows the doubling of the memory required as the state of each elemental filter is propagated to the next event time in step (a). In Figure 9, the update phase, step (b), is shown. The elemental covariance matrices are propagated and updated in exactly the same manner.

The complexity of the expressions and the expanding number of individual filters strongly motivate the simplifications to the estimator discussed in Chapter V.

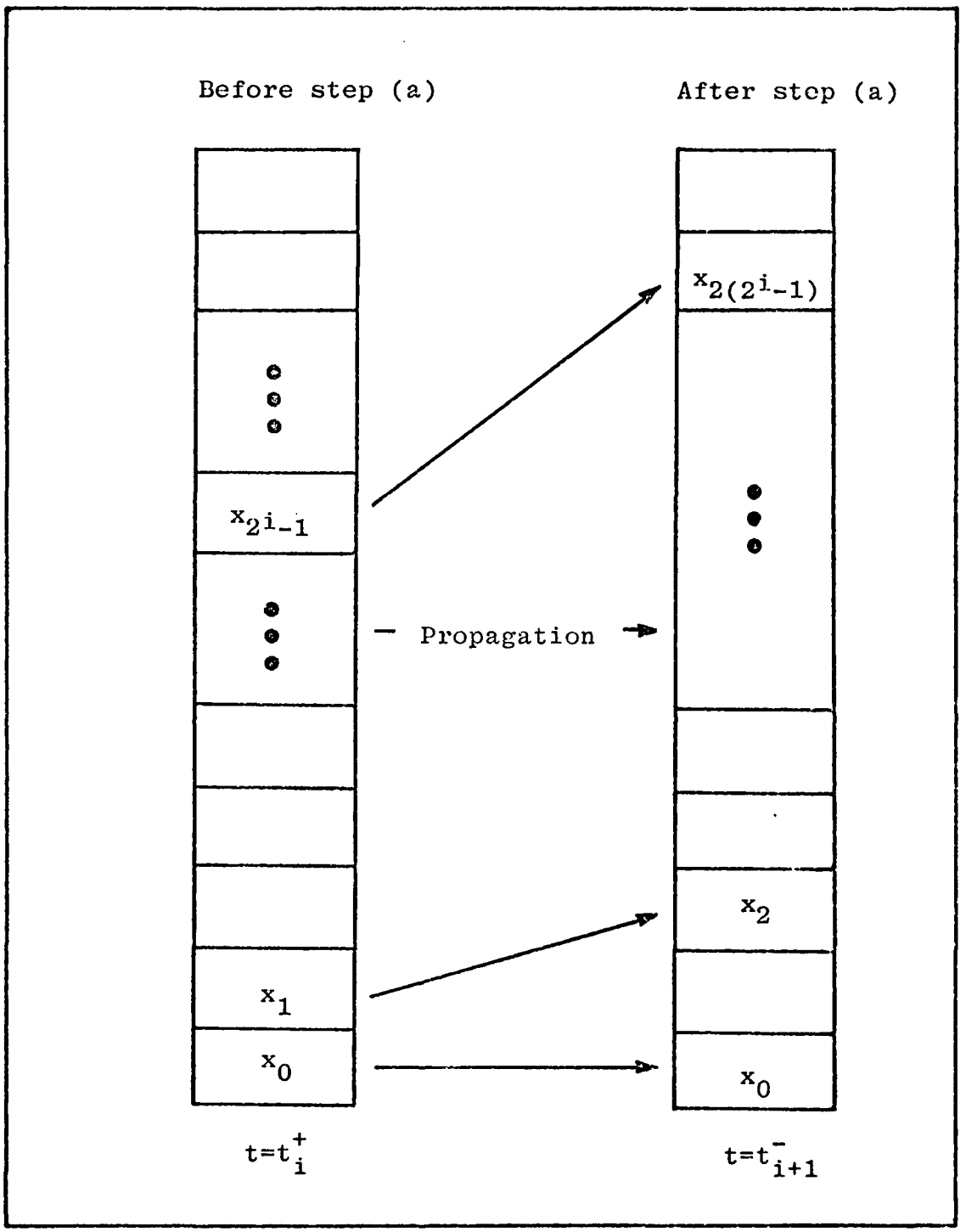


Figure 8. $\bar{x}$ storage before and after algorithm step (a).

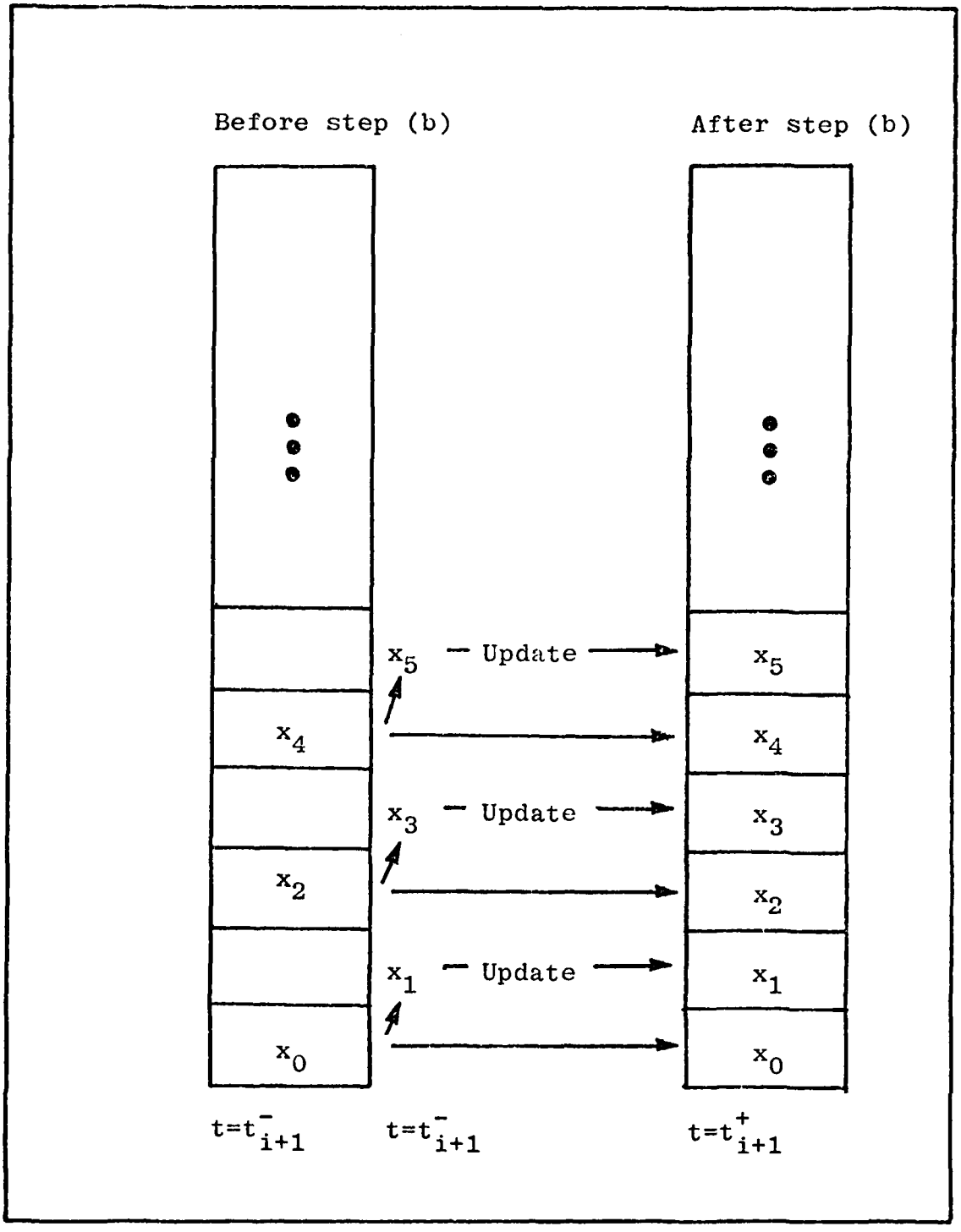


Figure 9. $\bar{x}$ storage before and after algorithm step (b).

IV.5 Summary

In this chapter, we have developed the weighting factors, $Pr[h_j^{N_t} | \mathcal{B}_t]$, for the multiple model adaptive estimator presented in Chapter II. The development is possible because of the cross product space modeling concepts and results of Chapter III, particularly the representation theorem. The multiple model adaptive estimator equations are consolidated in this chapter and several examples are given to provide insight into the estimator's structure and calculations. Simplifications to the estimator are discussed in the next chapter.

V. Filter Simplifications

V.1 Introduction

The multiple model adaptive estimator developed in the previous chapters provides the minimum mean square error estimate of a vector $\bar{x}$ (which is not directly observable) from observations of a point process signal in additive, independent point process noise. The estimator consists of elemental filters, the number of which grows exponentially with each new observed space-time event. In addition, the calculation of the $\hat{\lambda}_s$ and $\hat{\lambda}_n$ terms of the weighting probabilities (equation (175)) is, in general, complex. These two factors motivate the simplifications developed here.

In this chapter, we develop simplifications to the multiple model adaptive estimator which require less computation than the methods of the full scale estimator. The simplified filter does not have a growing requirement for either memory or computation as additional space-time events are observed. The simplifications result in a suboptimal filter; however, they are based on an understanding of the structure of the full scale estimator and a brief discussion is given of the conditions under

which these simplifications are appropriate.

Two areas are considered: calculation of the probability weighting factors, and limitation of the exponential growth of the number of elemental filters. We begin by considering calculation of the weighting factors.

V.2 Weighting Factor Simplification

The form of the general weighting factors as, shown in equation (175), is very straightforward. The key to the complexity (or simplicity) of calculation is in the evaluation of the terms $\hat{\phi}_s$ and $\hat{\phi}_n$.

The term $\hat{\phi}_s$ is defined as

$$\hat{\phi}_s(t,\bar{r};\omega) = E_s\{\phi_s(t,\bar{r};\omega;\omega_s)|\mathcal{B}_t\} \tag{220}$$

and $\hat{\phi}_n$ is defined in a similar manner. The difficulty in calculating these two terms is highly dependent on the model of the underlying physical process. At one extreme, if we make the assumption that observed process is Poisson with known signal rate parameter

$$\phi_s(t,\bar{r};\omega;\omega_s) \triangleq \lambda_s(t,\bar{r},\bar{x}(t)) \tag{221}$$

and if $\bar{x}(t)$ is known, then

$$\hat{\phi}(t,\bar{r};\omega) = \lambda_s(t,\bar{r},\bar{x}(t)) \tag{222}$$

because ϕ_s is not dependent on ω_s. At the other extreme, we could, conceivably, formulate a model for which there is no closed form solution for $\hat{\phi}_s$ and $\hat{\phi}_n$.

A more useful case is that of the particle beam problem where λ_s is defined by equation (3) and λ_n is a deterministic function of t and $\bar{r}$. (A random noise rate parameter is allowed and useful for the beam problem. We are given the form of λ_s and there are significant difficulties in computing $\hat{\lambda}_s$. Because we have considerable freedom in the form of λ_n, only simplifications to the $\hat{\lambda}_s$ calculation are considered here.)

Additionally, let there be no feedback for this example. Since λ_n is not random

$$\hat{\lambda}_n(t,\bar{r}) = \lambda_n(t,\bar{r}) \tag{223}$$

and we must evaluate

$$\hat{\lambda}_s(t,\bar{r}) = \int_{R^n} \lambda_s(t,\bar{r},\bar{\xi}) f(\bar{\xi}|Z^{N_t}) d\bar{\xi} \tag{224}$$

where $\bar{\xi}$ is the dummy variable for $\bar{x}(t)$ and $f(\xi|Z^{N_t})$ is the probability density function of $\bar{x}(t)$ given the

measurement history. We can use Bayes' rule to obtain

$$f(\bar{\xi}|Z^{N_t}) = \frac{f(Z^{N_t}|\bar{\xi})f(\bar{\xi})}{f(Z^{N_t})} \tag{225}$$

in which $f(Z^{N_t})$ is the sample function density of the observed process and $f(\bar{\xi})$ is the probability density function of $\underset{\sim}{\bar{x}}(t)$ propagated from time t_0 .

From the definition of a sample function density, we can write

$$\frac{f(Z^{N_t}|\bar{\xi}_g)}{f(Z^{N_t})} = \exp\left[-\int_{t_0}^{t}\int_{Y}\phi(\tau,\bar{\pi},\bar{\xi}_g)-\hat{\phi}(\tau,\bar{\pi})d\bar{\pi}d\tau\right.$$

$$\left. + \int_{t_0}^{t}\int_{Y}\ln\left[\frac{\phi(\tau,\bar{\pi},\bar{\xi}_g)}{\hat{\phi}(\tau,\bar{\pi})}\right]N(d\tau \times d\bar{\pi})\right] \tag{226}$$

where $\phi \triangleq \lambda_s(t,\bar{r},\bar{\xi}) + \lambda_n(t,\bar{r})$ and $\bar{\xi}_g$ is the given value of $\underset{\sim}{\bar{x}}(t)$ in the numerator probability density function in equation (225).

We now make the assumption that $Y=R^m$. To satisfy the

regularity conditions, $\lambda_n(t,\bar{r})$ must now be integrable over R^m. This assumption, that $Y=R^m$, is not too restrictive, however, because we can usually redefine $\lambda_n(t,\bar{r})$ to be integrable over some region of interest and zero elsewhere.

As a result of this assumption, we can simplify equation (226) further. Consider the term

$$\int_{t_o}^{t}\int_{R^m}(\phi-\hat{\phi})d\bar{\pi}d\tau \triangleq \int_{t_0}^{t}\int_{R^m}\phi(\tau,\bar{\pi},\bar{\xi}_g)-\hat{\phi}(\tau,\bar{\pi})d\bar{\pi}d\tau \tag{227}$$

$$= \int_{t_0}^{t}\int_{R^m}\phi(\tau,\bar{\pi},\bar{\xi}_g)-E_s\{\phi(\tau,\bar{\pi},\bar{\xi})|Z^{N_t}\}d\bar{\pi}d\tau$$

We can integrate the terms separately and interchange the order of integration and expectation by the Fubini theorem to obtain

$$\int_{t_0}^{t}\int_{R^m}\phi-\hat{\phi}\,d\bar{\pi}d\tau$$

$$=\int_{t_0}^{t}\left[\int_{R^m}\phi(\tau,\bar{\pi},\bar{\xi}_g)d\bar{\pi}-E_s\left\{\int_{R^m}\phi(\tau,\bar{\pi},\bar{\xi})d\bar{\pi}\,|\,Z^{N_t}\right\}\right]d\tau \tag{228}$$

Because the assumed Gaussian shape of $\lambda_s(t,\bar{r},\bar{\xi})$ results in a tractable integration, we can evaluate the integral

$$\int_{R^m}\phi(t,\bar{\pi},\bar{\xi})d\bar{\pi}$$

$$=\alpha_1(t;\omega_{s_3})\int_{R^m}\lambda_s(t,\bar{\pi},\bar{\xi})d\bar{\pi}+\alpha_0(t;\omega_{s_3})\int_{R^m}\lambda_n(t,\bar{\pi})d\bar{\pi} \tag{229}$$

$$=\Lambda(t)(2\pi)^{\frac{m}{2}}|\underline{R}(t)|^{\frac{1}{2}}+\int_{R^m}\lambda_n(t,\bar{\pi})d\bar{\pi} \tag{230}$$

where the last step follows from equation (120) and the fact that no conditioning on a particular hypothesis sequence is given, therefore

$$\alpha_0(t;\omega_{s_3}) = \alpha_1(t;\omega_{s_3}) = 1 \tag{231}$$

Note that $\Lambda(t)(2\pi)^{\frac{m}{2}}|\underline{R}(t)|^{\frac{1}{2}}$ is not a function of ω or ω_s, even if feedback control is included in the manner previously described. For the assumed λ_n

$$\int_{R^m} \lambda_n(t,\bar{\pi})d\bar{\pi} = E_s\{\int_{R^m} \lambda_n(t,\bar{\pi})d\bar{\pi}|Z^{N_t}\} \tag{232}$$

therefore

$$\int_{t_0}^{t}\int_{R^m} \phi-\hat{\phi}\ d\bar{\pi}\ d\tau = 0 \tag{233}$$

and equation (226) reduces to

$$\frac{f(Z^{N_t}|\bar{\xi})}{f(Z^{N_t})} = \exp\left[\int_{t_0}^{t}\int_{R^m} \ln\left[\frac{\phi(\tau,\bar{\pi},\bar{\xi}_g)}{\phi(\tau,\bar{\pi})}\right] N(d\tau \times d\bar{\pi})\right] \tag{234}$$

From the definition of a counting integral, we can write equation (234) as

$$\frac{f(Z^{N_t}|\bar{\xi})}{f(Z^{N_t})} = \prod_{i=1}^{N_t} \frac{\phi(t_i,\bar{r}_i,\bar{\xi}_g)}{\hat{\phi}(t_i,\bar{r}_i)} \tag{235}$$

where the index, i , corresponds to the observed point process events $\{\bar{z}(1), \bar{z}(2), \ldots, \bar{z}(i), \ldots \bar{z}(N_t)\}$.

By substituting equation (235) into equation (225) and equation (224), we have

$$\hat{\lambda}_s(t,\bar{r}) \prod_{i=1}^{N_t} \hat{\phi}(t_i,\bar{r}_i)$$

$$= \int_{R^m} \lambda_s(t,\bar{r},\bar{\xi}) \prod_{i=1}^{N_t} \phi(t_i,\bar{r}_i,\bar{\xi}) f(\bar{\xi}) d\bar{\xi} \tag{236}$$

Equation (236) is, in principle, solvable, however it requires an increasing amount of calculation as more point process events are observed. The order of the polynomial in $\hat{\lambda}_s$ (which must be solved at each event time) is N_t+1 and the product of ϕ terms in the integral increases by one at each observation time. In addition, $f(\bar{\xi})$ is the Gaussian probability density function for $\bar{\underset{\sim}{x}}(t)$. The covariance of this density increases substantially as the density is propagated forward in time. This will eventually lead to numerical problems in an implementation of the estimator.

Due to these difficulties in evaluating $\hat{\lambda}_s$ for this example, consider the following simplification:

> To evaluate $\hat{\lambda}_s(t_i,\bar{r}_i)$ take the estimate, $\hat{\bar{\underset{\sim}{x}}}(t_{i-1}^+)$ of $\bar{\underset{\sim}{x}}(t_{i-1})$ and propagate it forward in time to t_i. Use this propagated estimate, $\hat{\bar{\underset{\sim}{x}}}(t_i^-)$ to evaluate equation (3) to obtain a value for $\hat{\lambda}_s$.

Two points should be noted. First, since equation (3) is not linear in $\bar{\underset{\sim}{x}}(t)$ in general $E\{\lambda_s(\bar{\underset{\sim}{x}}(t))\}$ will not be exactly equal to $\lambda_s(E\{\bar{\underset{\sim}{x}}(t)\})$, and it is the latter form we are using in this simplification. Second, since we are propagating an old estimate of $\bar{\underset{\sim}{x}}(t_{i-1})$ to use at time t_i , it is important that we use an accurate propagation model. If the assumed model of the underlying process (equation (4)) has characteristics different than those of

the actual physical process, then we would expect this simplification to result in poor performance due to the inaccurate propagation. We have assumed perfect knowledge of the dynamics of the underlying physical process throughout this research. Thus, this simplification is reasonable.

V.3 Limitation to a Fixed Number of Elemental Filters

In this section, we consider a method of limiting the exponential growth of the full scale multiple model adaptive estimator so that (after the startup of the filter) there is always a fixed number of elemental filters. The overall method is to consider only observations from a finite data window.

Let D be the "depth" of the algorithm where the depth is the number of the most recent point process observed events which are explicitly included in the calculation of the elemental filter estimates. If D events are considered, then there are 2^D elemental filters and the number of filters is constant. As each new point process event is observed, we must eliminate the "oldest" observed event and incorporate the new event. Three methods for accomplishing this are discussed in the following sections.

V.3.1 Method One: Re-initialization. The first method, called re-initialization, restarts the entire multiple model adaptive estimator each time a new point process event is observed, which would result in more than 2^D elemental filters. The new event is added to the set of observations considered in the algorithm and the oldest observed event is discarded so that there is a maximum of D observed events explicitly present in the MMAE calculation.

The algorithm for the re-initialization method is as follows. Recall that N_t is the number of point process events in the interval $[t_0,t)$.

(a) For $N_{t_i} \leq D$:

Operate as a full scale MMAE filter as described in Section IV.3.

(b) For $N_{t_i} > D$:

(1) Re-initialize the filter by letting $\hat{\underline{x}}(t_{i-D})$ be the new initial condition at time t_{i-D} .

(2) Propagate and update the elemental filters from time t_{i-D} to time t_i .

(3) Calculate $\hat{\underline{x}}(t_i)$ as a weighted sum as described in Section IV.3.

Even though only the most recent D observations are explicitly included in the calculation, the older observed events influence the value of $\hat{\bar{x}}(t_{i-D})$. Thus we fix the number of elemental filters by a data window concept but the entire measurement history is still considered in the calculation of $\hat{\bar{x}}(t_i)$.

A major disadvantage of this method is that the entire set of 2^D elemental filters must be propagated over D inter-event intervals and the updates due to D events must be incorporated for each newly observed point process event.

V.3.2 Method Two: Strict Window. The second method, termed "strict window", completely disregards all observed events except the most recent D events. This is accomplished by assuming that all events prior to the most recent D events were caused by noise. The algorithm for the strict window method is as follows:

(a) For $N_{t_i} \leq D$:

Operate as a full scale MMAE filter as described in Section IV.3.

(b) For $N_{t_i} > D$:

(1) Retain only the $(2^D)/2$ elemental filters from the lower half of the hypothesis sequence tree at time t_{i-1} .

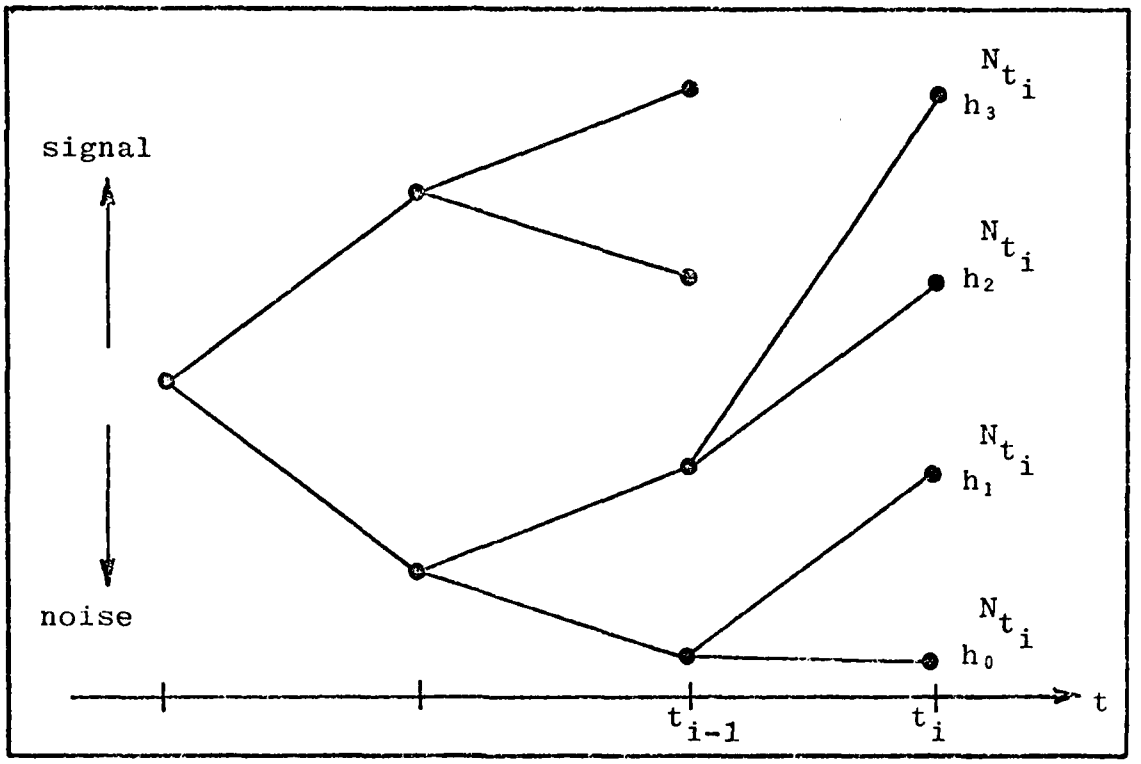


Figure 10. Strict Window Method

(2) Propagate the $(2^D)/2$ filters to time t_i and update them to obtain 2^D elemental filters.

(3) Calculate $\hat{\bar{x}}(t_i)$ as in Section IV.3.

The strict window method is depicted in Figure 10 for the sample case of D=2. The bottom hypothesis sequence, h_0, is that which assumes all observed events are due to noise. Because of this, we do not need to re-initialize, repropagate, or re-update all 2^D elemental filters as each new event is observed; the calculations have already been

performed.

A major disadvantage of the strict window method is that the "old" observed point process events which are outside the strict window are completely ignored. They have no effect on the value of $\hat{\bar{x}}(t_i)$ and thus we are not gaining information about $\bar{x}(t_i)$ from the measurement history from time t_0 to time t_{i-D} . A second disadvantage of the Strict Window method is that it relies completely on the filter's internal model of the dynamics of the underlying process for the time interval prior to the data window. This could cause serious errors if the filter's model were not correct.

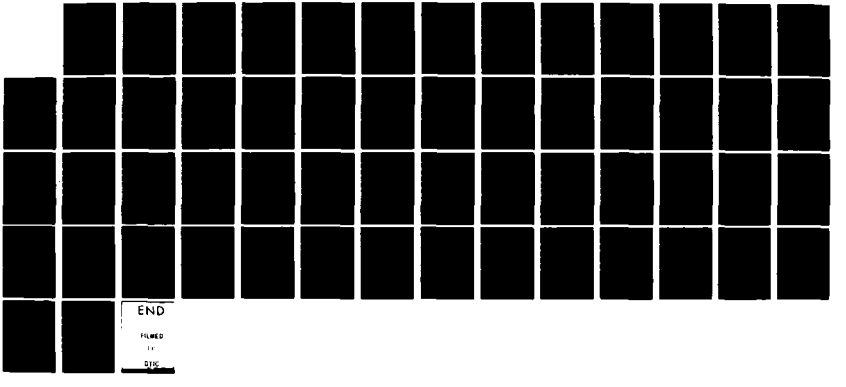
V.3.3 Method Three: Best Half. The third method, termed best half, fixes the number of elemental filters, uses information about $\bar{x}(t_i)$ from the entire measurement history as in method one, and retains the efficiency of calculation as in method two. The best half algorithm is as follows:

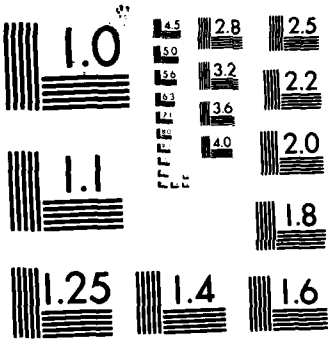
(a) For $N_{t_i} \leq D$:

Operate as a full scale MMAE filter as described in Section IV.3.

AD-A124 827 MULTIPLE MODEL ADAPTIVE ESTIMATION FOR SPACE-TIME POINT PROCESS OBSERVATIONS(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL OF ENGINEERING D E MEER 3/3

UNCLASSIFIED AUG 82 AFIT/DS/EE/82-2 F/G 12/1 NL





(b) For $N_{t_i} > D$:

(1) Calculate:

$$\Pr[\text{upper half}] = \sum_{j=\frac{2^D}{2}}^{2^D-1} \Pr[h_j^{N_{t_i}} | \mathcal{B}_t]$$

and

$$\Pr[\text{lower half}] = \sum_{j=0}^{\frac{2^D}{2}-1} \Pr[h_j^{N_{t_i}} | \mathcal{B}_t]$$

(2) If: Pr[upper half] > [Pr lower half]

Propagate the upper half of the hypothesis sequences to time t_i and update them with the observation at time t_i .

Otherwise:

Propagate and update the lower half.

The best half algorithm is shown in Figure 11 for the case of D=2. The solid lines depict the propagation and update paths when the upper half is the most probable set of hypotheses. The dashed lines depict the propagation and update paths when the lower half is the most probable set of

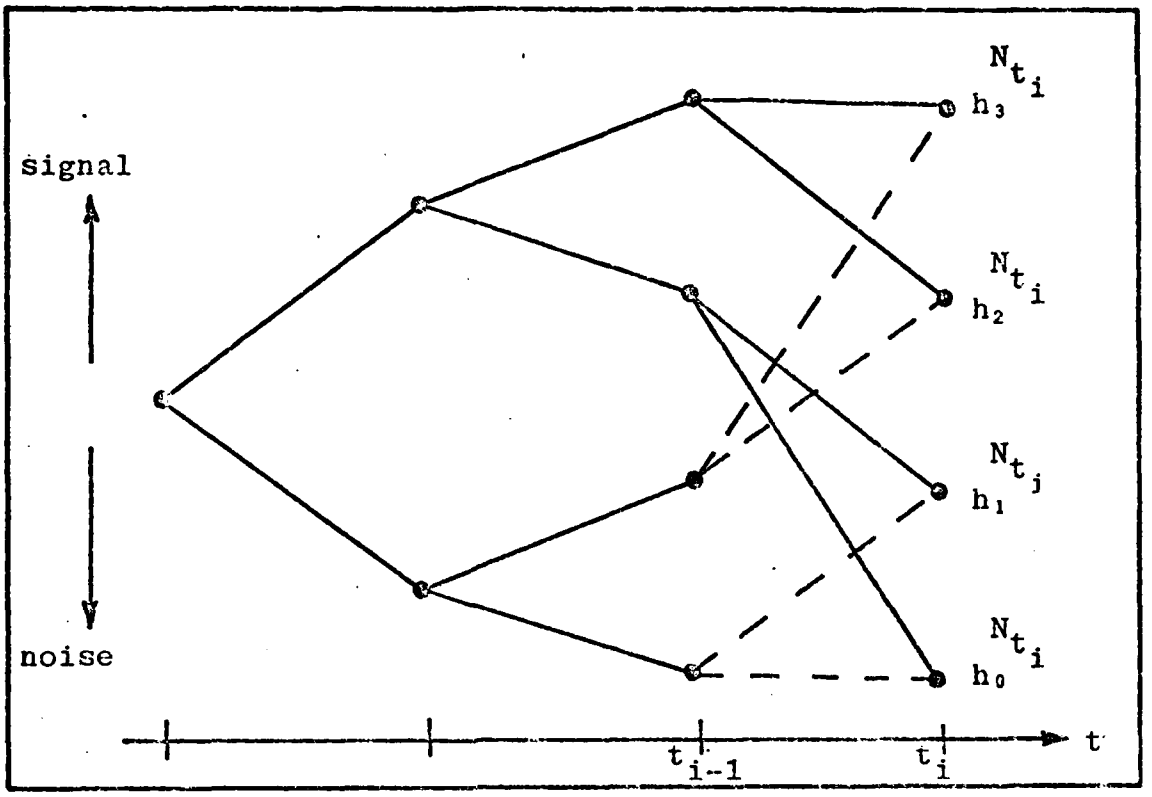


Figure 11. Best Half Method

hypotheses. Note that if the lower half is chosen as most probable, then the propagation and update steps for that observation are identical to those of the strict window method. In essence, the best half algorithm makes a final decision at time t_i as to whether the event at time t_{i-D} was caused by signal or noise. Because the algorithm can select from either half of the set of hypotheses, the entire measurement history influences $\hat{\bar{x}}(t_i)$. In addition, since the algorithm selectively retains half of the possible hypotheses, it is not necessary to re-initialize, repropagate, or re-update all 2^D elemental filter over D

inter-event intervals each time a new event is observed.

Due to its efficiency of calculation and consideration of the entire measurement history, the best half method is recommended as the best means (for the three methods considered) of limiting the estimator to a fixed number of elemental filters.

V.4 Summary

Two major areas of simplification for the MMAE algorithm are addressed in this chapter. First, the complexity of evaluating the weighting factors is reduced, in Section V.2, by approximating λ_s with an estimate of λ_s based on the previous observed point process events. This simplification is reasonable due to the assumed Poisson statistics of the signal process and the assumed perfect knowledge of the dynamics of the underlying process. Second, the exponential growth in memory and calculation required is avoided by using data window concepts to keep the number of elemental filters fixed. Three possible methods of achieving a fixed number of filter are discussed and the "best half" algorithm (Section V.3.3) is recommended due to its relative efficiency of calculation and incorporation of the entire measurement history.

Results of Monte Carlo simulations of a multiple model adaptive estimator using these simplifications are presented in Chapter VI.

VI. Simulation Results

VI. Introduction

The results of a Monte Carlo simulation of the simplified multiple model adaptive estimator are presented in this chapter. All of the results are from an estimator implemented using the approximation for calculation of (Section V.2) and the Best Half method for limiting the estimator to a fixed number of elemental filters (Section V.3.3). These simplifications were made to ease the large computation and storage requirements of the full-scale estimator, as discussed in Chapter V.

The goal of these simulations is to determine the trends in the error of the estimator as several of the major parameters are varied and also to investigate the sensitivity to these parameters. Although the actual values used for the parameters are appropriate for a tracking type application, they are not taken from an actual tracking problem. Therefore, the error performance results are useful as a relative measure of performance as various parameters are changed rather than as an absolute measure of the estimator's performance for a particular application.

In Section VI.2, the specific model used in this

simulation is described. Section VI.3 contains some preliminary results including sample track plots and the sensitivity of the estimator to the number of runs in the Monte Carlo simulation. In Section VI.4.1 through VI.4.6, the sensitivity of the estimator to six major parameters is shown. Section VI.4.7 contains results of several simulations to test the ability of the estimator to acquire the true value given poor initial conditions. An initial lower and upper bound on the performance of the estimator are given displayed in Section VI.4.8. Some other simulation considerations are discussed in Section VI.5 including the effects of mismatched models and two suggested methods of dealing with the mismatch.

VI.2 The Simulation Model

The model described here is of a one dimensional tracking application in which there is no feedback control. Let λ_s be defined as in equation (3) where

$$\begin{aligned} m &= 1 \\ \Lambda(t) &= 1 \\ \underline{H}(t) &= 1 \\ \underline{R}(t) &= R \text{ (scalar)} \end{aligned} \tag{237}$$

Let the process $\underset{\sim}{\bar{x}}(t)$ be defined in differential form by

equation (4) where

$$\begin{aligned} \underset{\sim}{\bar{x}}(t) &= x(t)\epsilon R^n \\ n &= 1 \\ \underline{G}(t) &= g \quad \text{(scalar)} \\ \underline{F}(t) &= -\frac{1}{\tau} \end{aligned} \tag{238}$$

and where $\underset{\sim}{\bar{u}}(t)$ is a one dimensional Wiener Process of unit diffusion.

Let the photo-electron event detector be modeled as a 10 centimeter (cm) interval on the real line centered about zero. Thus r and x have dimensions of centimeters and the dimension of R is centimeters squared.

We define the noise rate parameter as

$$\lambda_n(t,\bar{r};\omega_s) = \begin{cases} \lambda_n & -5\text{cm} \leq r \leq 5\text{cm} \\ 0 & \text{elsewhere} \end{cases} \tag{239}$$

This form is selected to model point process noise events caused by a uniform dark current mechanism in a continuous detector over the interval $-5\text{cm} \leq r \leq 5\text{cm}$. A simple random $\underset{\sim}{\lambda}_n$ case is considered in Section VI.4.1.

All of the simulations are made over a 100 second time interval. For this one dimensional case, the expected number of signal events is proportional to the area under the Gaussian shaped λ_s function and the expected number of noise events is proportional to the area under the λ_n function defined by equation (239). In the simulation data that follows, the signal to noise ration is defined as

$$SNR = \frac{\Lambda\sqrt{2\pi R}}{\lambda_n L} \tag{240}$$

where L is the length of the detector. The numerator of equation (240) is the expected number of signal point process events per unit time interval and the denominator is the expected number of noise point process events per unit time interval. Thus, the SNR for this point process model is the ratio of the expected number of signal events to the expected number of noise events.

VI.3 General Results

Figure 12 shows the true value of x and the output of the multiple model adaptive estimator for one 100 second simulation of the filter. The true value of x is displayed by the solid line and the broken line is the

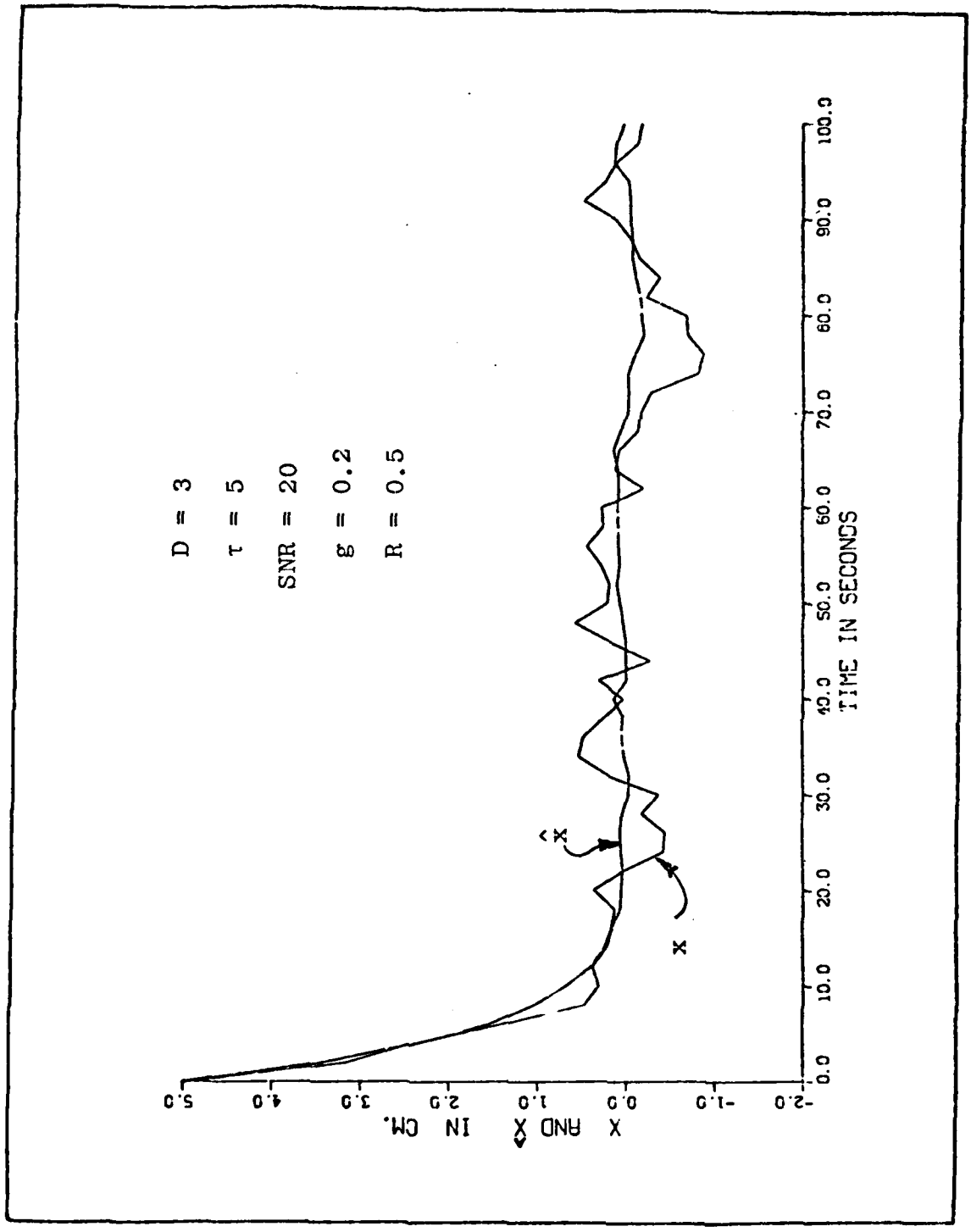


Figure 12. Example of True and Estimated x(t) Values for One Run.

filter's estimate of x. The values of the various parameters are listed on the figure. In all of these figures, the expected number of signal point process events is 100 unless noted otherwise (as in Section VI.4.2). This sample run was made by passing to the filter the true initial conditions, x_0=5cm. The smoothness of $\hat{x}$ with respect to x is inherent because $\hat{x}$ is a weighted sum of up to eight elemental filter estimates, and because the filter is given the exact dynamical model for x.

Figure 13 shows the ensemble averages over 50 simulation runs of x and $\hat{x}$ for the same set of parameters and initial conditions displayed in Figure 12. The ensemble averages of the error statistics, for this 50 run example, are shown in Figure 14. The solid line in Figure 14 is the ensemble average of the error, where the error is defined as $x-\hat{x}$ The two irregular broken lines are the ensemble averages of the error plus or minus the standard deviation of the error. The two relatively smooth dashed lines are plus and minus the square root of the filter variance. The filter variance is the filter's estimate of how well it is performing. Of importance here is the fact that the filter's estimate of its error is similar to its actual performance and the filter variance neither diverges nor goes to zero.

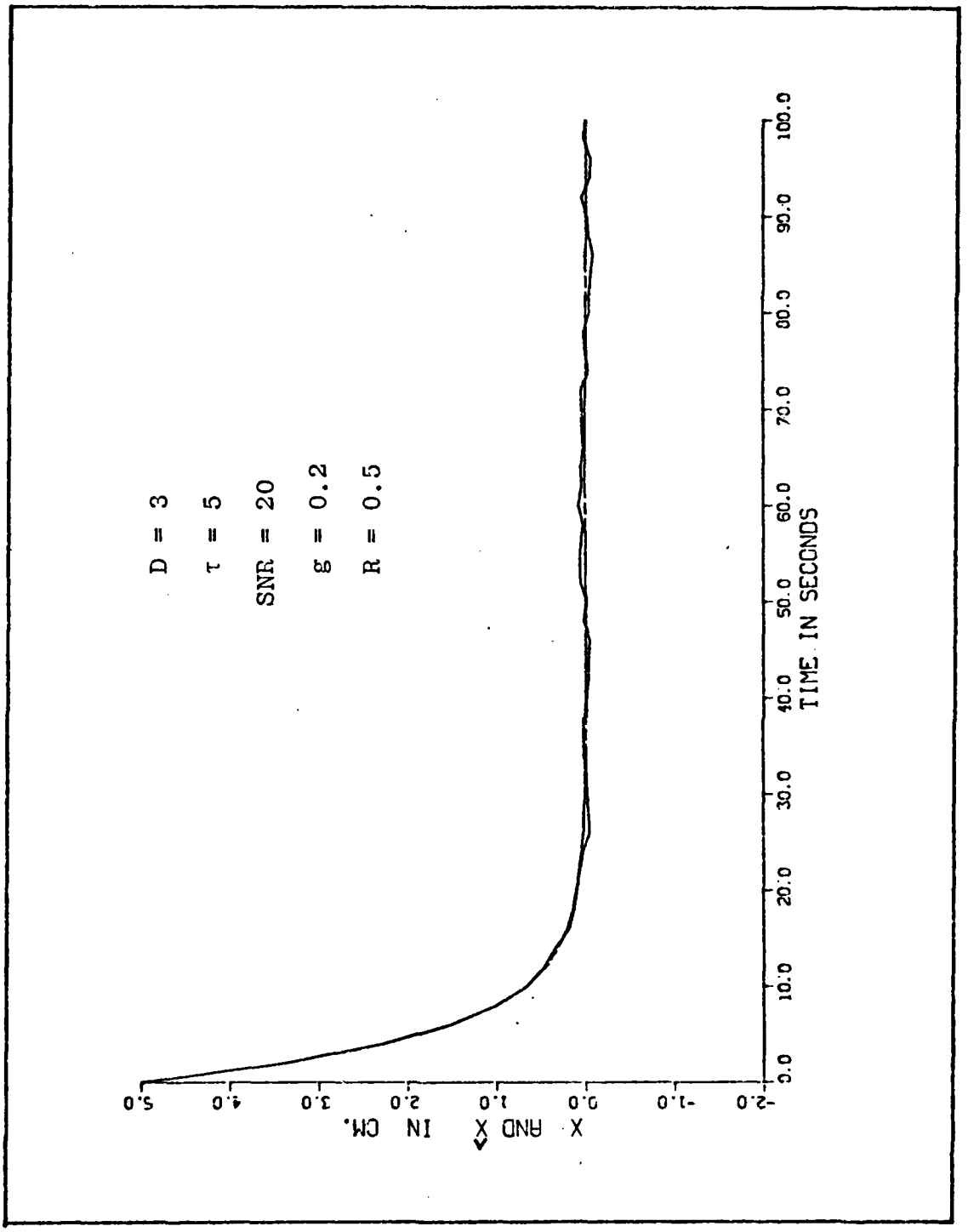


Figure 13. True and Estimated Ensemble Averages for 50 Runs.

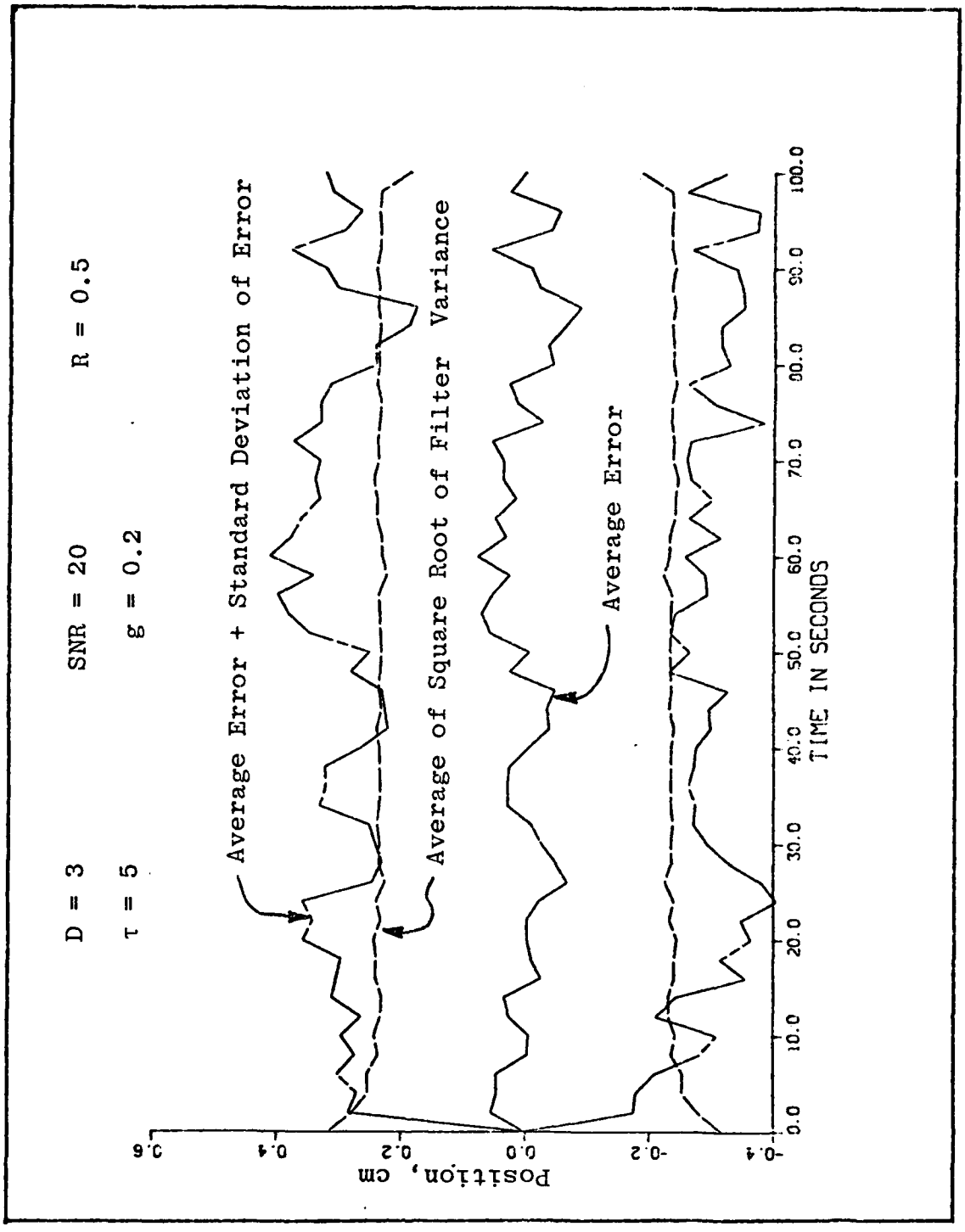


Figure 14. Ensemble Error Statistics for 50 Runs.

The smoothing of the data, as more simulation runs are made, is typical of Monte Carlo simulations. A pertinent question is, how many runs are sufficient to give accurate performance data without expending an excessive amount of computation time. One method of addressing this question is to vary the number of simulation runs for a fixed set of input parameters and observe the error statistics. As more runs are made, the error statistics should converge to a final value. The results of this analysis are shown in Figure 15. In this figure, the number of runs is varied from two to 100 and the root mean squared (RMS) error at time t=50 seconds is plotted. The RMS error is chosen as the measure of performance in this (and subsequent sensitivity tests) because it gives a measure of the error from the true value regardless of sign. If we used the absolute error, a positive error on one run could cancel a negative error on another run, resulting in an incorrectly low ensemble average. The time for sampling the error (t=50 seconds) is chosen to minimize the effect of the filter startup on the performance results. In Figure 15, we are not interested in the actual value of the RMS error. Instead, we are looking for a point beyond which there is little change in the error. The value of the error is erratic for numbers of runs less than 20. For 20 or more runs, there is little change in the final RMS error. Based

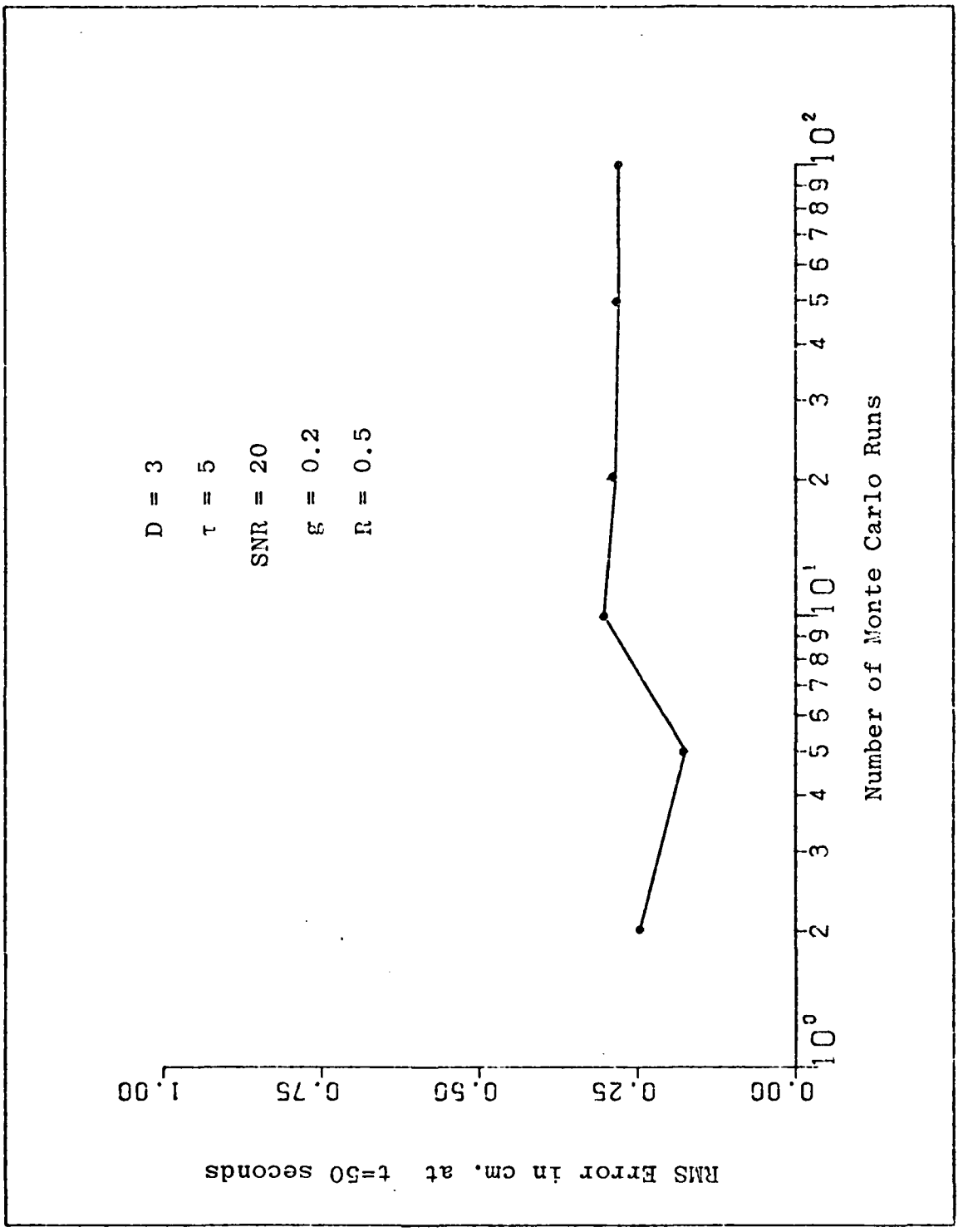


Figure 15. Sensitivity to Number of Monte Carlo Runs.

on this, the plotted results in all subsequent sections are for Monte Carlo simulations with ensemble averaging over 50 runs.

VI.4 Parameter Sensitivity Results

In the following sections, the sensitivity of the estimator to changes in several of its major parameters is investigated. The general method is to vary one parameter while keeping the rest constant. The performance measure for evaluation of the sensitivity is the RMS error at time t=50 seconds. (The acquisition results are measured at t=16 seconds as discussed in Section VI.4.7.) In all cases, the number of Monte Carlo runs is 50. We begin by looking at the effect of the noise strength.

VI.4.1 Sensitivity to g. The parameter g is the square root of the strength of the white Gaussian noise driving the dynamics of the unobserved process x. The lower curve in Figure 16 shows the performance of the estimator as g is varied from 0.01 to 1.0 cm. The trend is as expected; as the dynamics of x increase, the ability of the estimator to track the changes diminishes and the RMS error increases.

The upper curve in Figure 16 shows the effect on the error of a random $\underset{\sim}{\lambda}_n$. The value of λ_n in equation (239) is calculated from the SNR and the expected number of signal

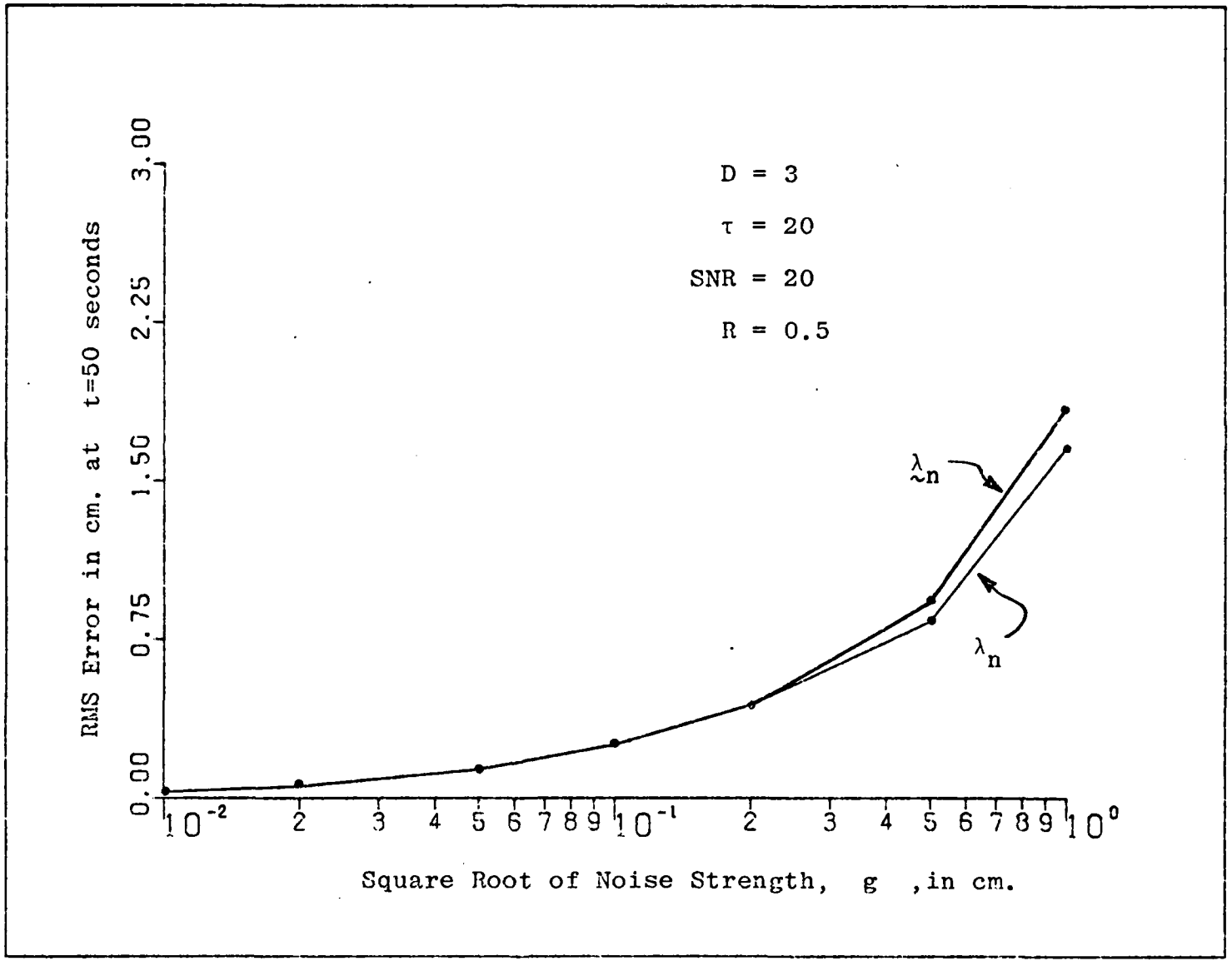


Figure 16. Estimator Performance Versus Dynamic Model Noise.

point process events. For the random $\underset{\sim}{\lambda}_n$ case, we let λ_n' be a uniform random variable on $[0,2\lambda_n]$ and a single realization of λ_n' is used for generation of the noise point process events for a single run of the random λ_n simulation. Note that the realization is used for the entire run and a new value is selected for the next run. The estimator uses the value λ_n for calculation of $\hat{\lambda}_n$. This procedure is chosen as a worst case condition; very little knowledge is assumed by the filter about the $\underset{\sim}{\lambda}_n$ process and no actual estimation of $\underset{\sim}{\lambda}_n$ is performed.

As can be seen from Figure 16, the RMS error increased for the random noise rate parameter case, but the increase is very small compared to the actual value of the error. For this set of parameters, the estimator is relatively insensitive to uncertainties in λ_n.

VI.4.2 Sensitivity to Expected Number of Signal Events. In Figure 17, the RMS error is plotted versus the expected number of point process signal events. As might be expected, as more information is available to the estimator from the signal process, the RMS error goes down. The trend is actually rather mild from signal counts ranging from one to 500. If the estimator only receives one signal-caused event in 100 seconds, then it must rely heavily on its

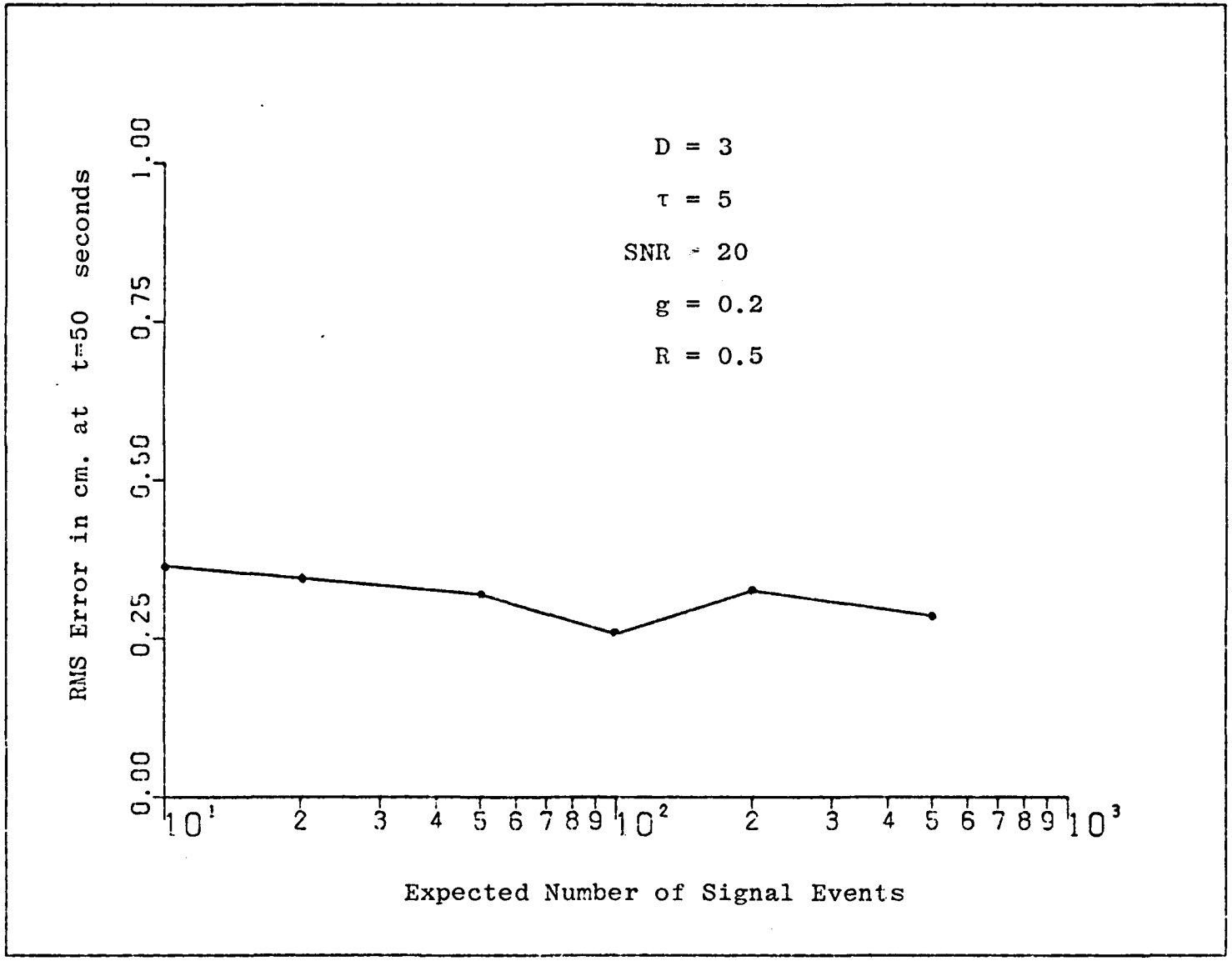


Figure 17. Estimator Performance versus Expected Number of Signal Events,

internal model for the dynamics of the x process. As described before, we have assumed perfect knowledge of the model. It is expected that the RMS error would be much larger at the low count rates if the dynamical models of the true process and the filter were mismatched.

VI.4.3 Sensitivity to τ. The sensitivity of the RMS error to τ is depicted in Figure 18. As can be seen, there is a strong trend toward increased RMS error as τ becomes larger. When τ is large, the dynamics of x depend proportionately more on the driving noise source and there is less restoring action due to the $\underline{F}(t)\bar{x}(t)dt$ term in equation (4). This allows errors between x and $\hat{x}$ to persist for longer periods between signal induced point process observation. When τ is small, any errors caused by the driving noise in equation (4) are rapidly reduced as the output decays to the steady state value.

VI.4.4 Sensitivity to SNR. Several SNR sensitivity tests were made for parameter sets similar to those shown in Figures 16 through 18. In all cases, there was virtually no effect on the RMS error, even for SNR values as low as 0.01.

The SNR sensitivity results shown in Figure 19 were obtained by setting τ and R to relatively large values (thus tending to raise the overall RMS error, see Figures 18 and 21) and by giving the multiple model adaptive estimator poor initial conditions. The trend displayed in Figure 19

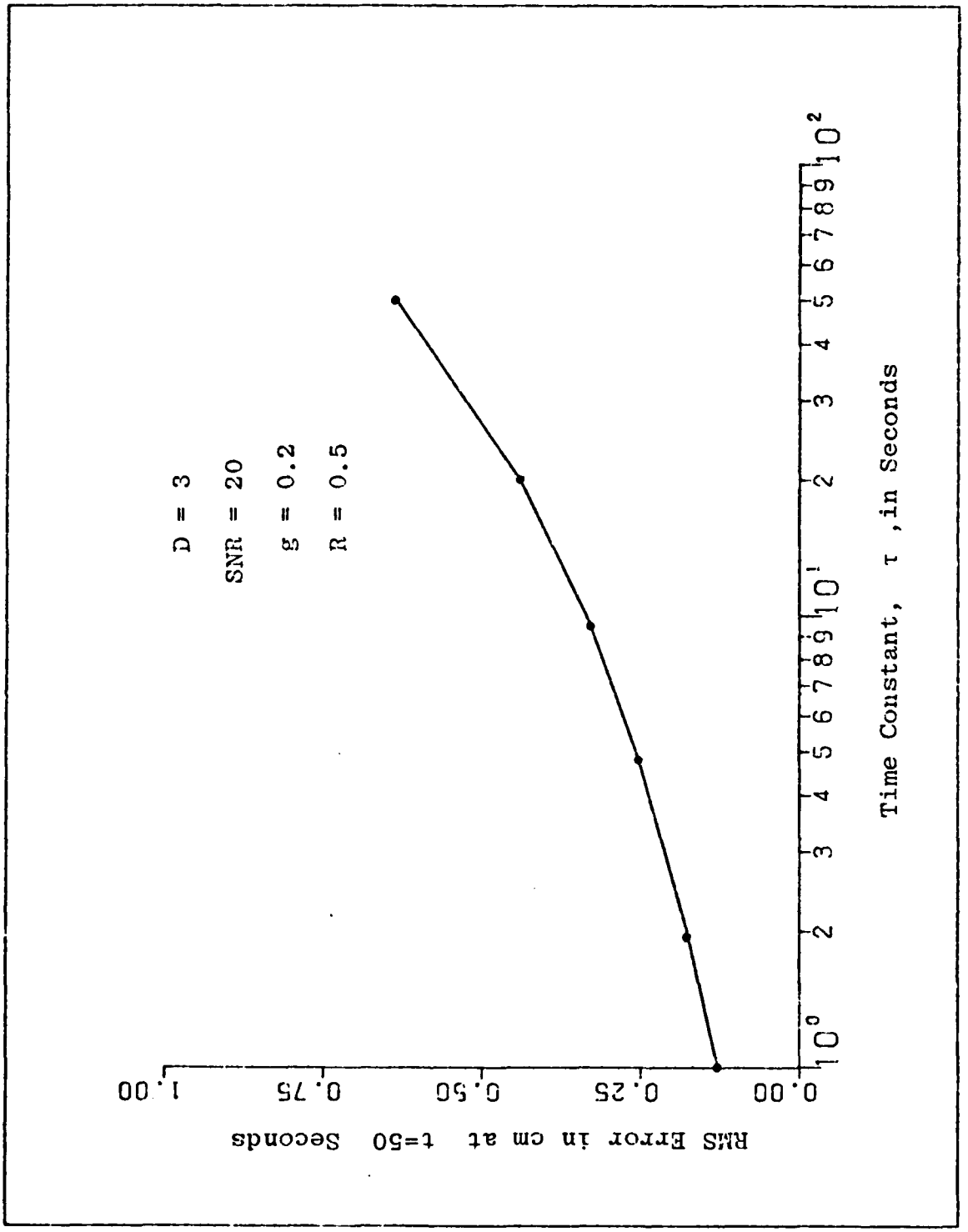


Figure 18. Estimator Performance versus τ .

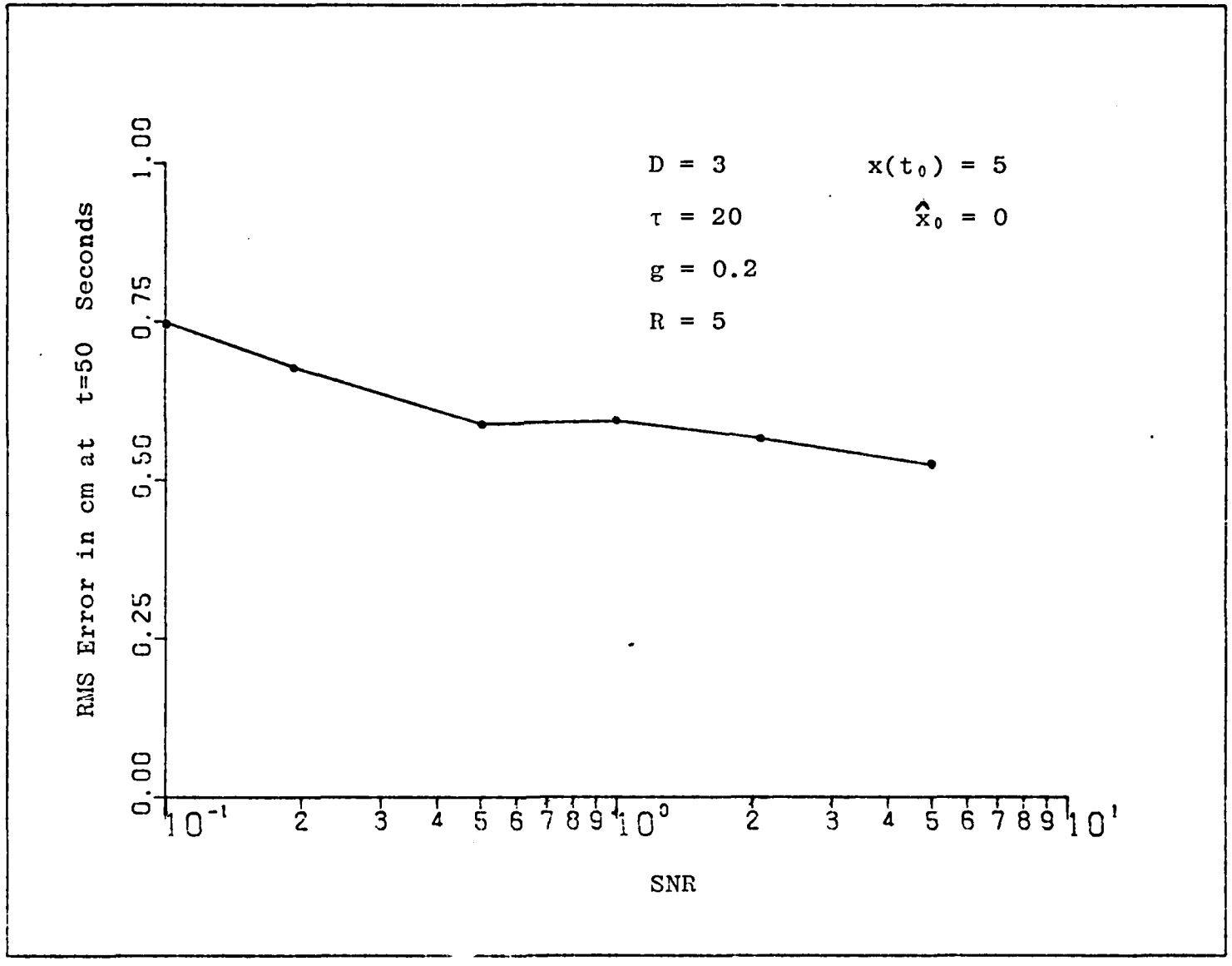


Figure 19. Estimator Performance versus Signal to Noise Count Ratio.

is as expected; the RMS error is lower for larger values of SNR. Of particular interest is the fact that the worst RMS error is only 0.75 cm for an expected signal to noise count ratio of 0.1 . That is, one signal event is expected for every 10 noise events. It is expected that the RMS error would increase more rapidly at low SNR values if the true and filter models were mismatched.

VI.4.5 Sensitivity to D. The sensitivity to depth, D , is shown in Figure 20. Note that the error axis scale on this plot is expanded and the variation in RMS error over a depth range of one to eight is very small. This suggests that if the model of the underlying process is well known then it may be possible to consider only the most recent observed event and obtain acceptable performance.

VI.4.6 Sensitivity to R. The sensitivity to R (the dispersion of the Gaussian shaped λ_s function) is displayed in Figure 21. The performance indicates that there are specific tuning considerations for R for a given set of parameter values. As R becomes very small, it appears that valid signal events are deweighted too heavily due to the narrow shape of λ_s. At the other extreme, when R is large, noise induced events near the signal source are not deweighted strongly enough. This characteristic curve, as R is varied, can also be seen in Figure 24.

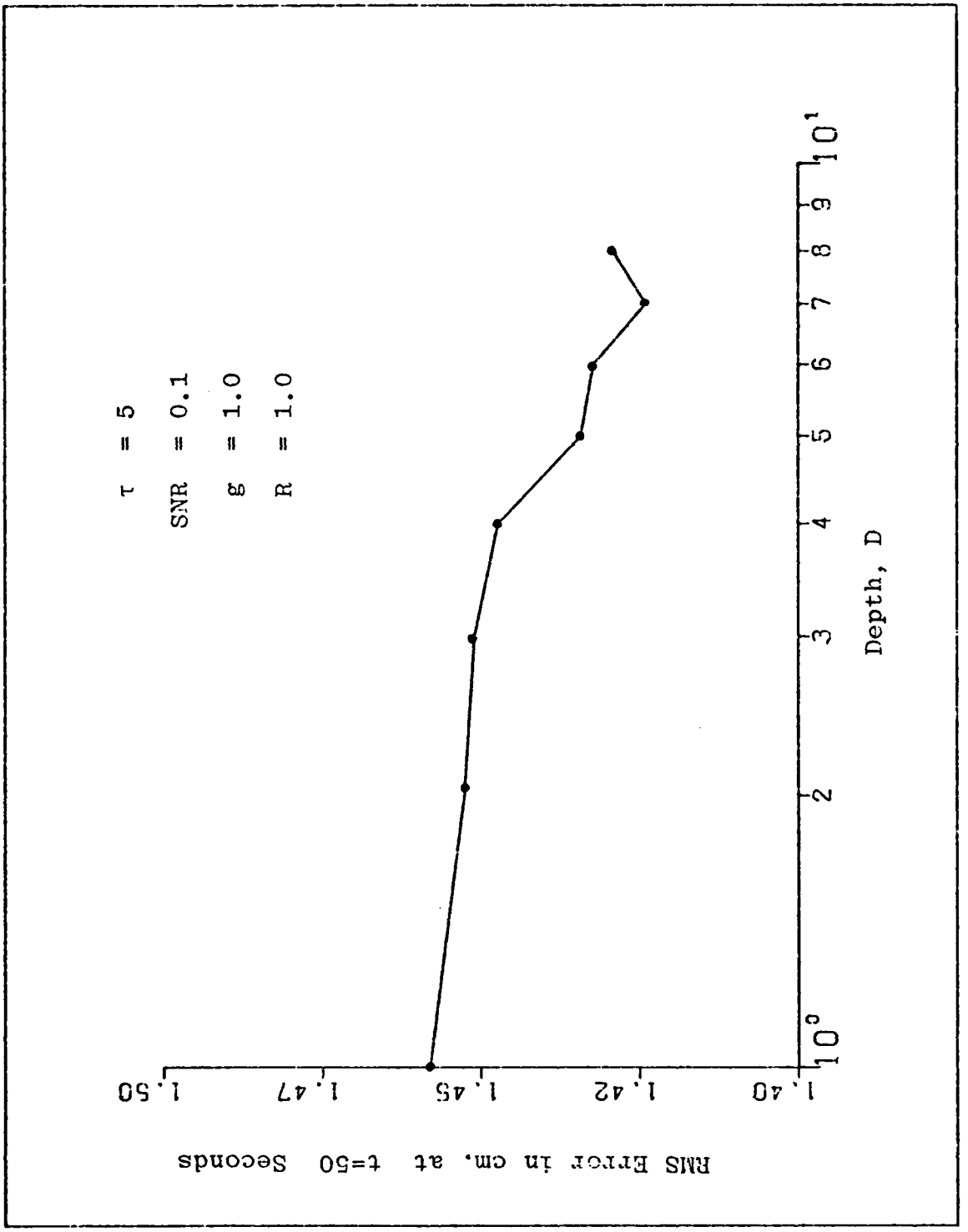


Figure 20. Estimator Performance versus Depth.

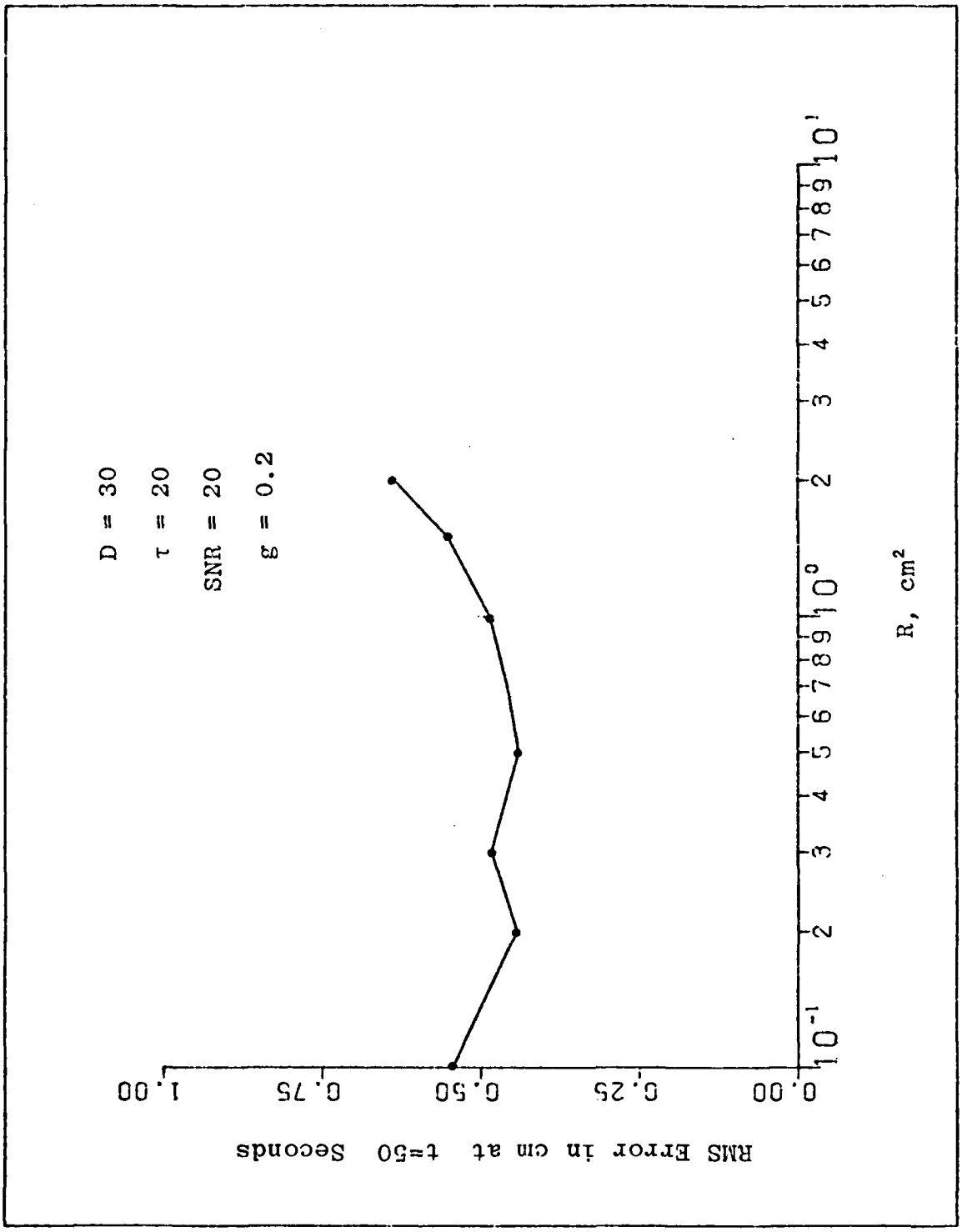


Figure 21. Estimator Performance versus R.

VI.4.7 Acquisition. A final set of parameters was selected to test the ability of the estimator to acquire the true x value given inaccurate initial conditions. In all of the following cases, $x_0=5$ and $\hat{x}_0=0$.

Figure 22 shows the ensemble average of x and $\hat{x}$ for the relatively small values of R =0.2 and $\hat{\Sigma}_0 = 0.1$. The performance is very poor until the true x value decays to a region close to $\hat{x}$. The acquisition can be greatly improved by setting R=2 and $\hat{\Sigma}_0=10$. These results are shown in Figure 23. Under the new conditions, the estimator quickly acquires the true x value and tracks it.

The RMS error performance versus R for these two values of $\hat{\Sigma}_0$ is displayed in Figure 24. Note that the values plotted are for t=16 seconds to insure that error in the acquisition region is being measured. Both curves show the characteristic tuning sensitivity to R as in Figure 21. The upper curve is the performance when the multiple model adaptive estimator has a high confidence in its initial conditions. The lower curve is the performance for the low confidence case.

VI.4.8 Performance Bounds. The estimator's RMS error versus square root of noise strength is plotted in Figure 25 along with an upper and lower bound on the RMS error.

The lower bound on performance was obtained by operating the estimator with only signal observed events,

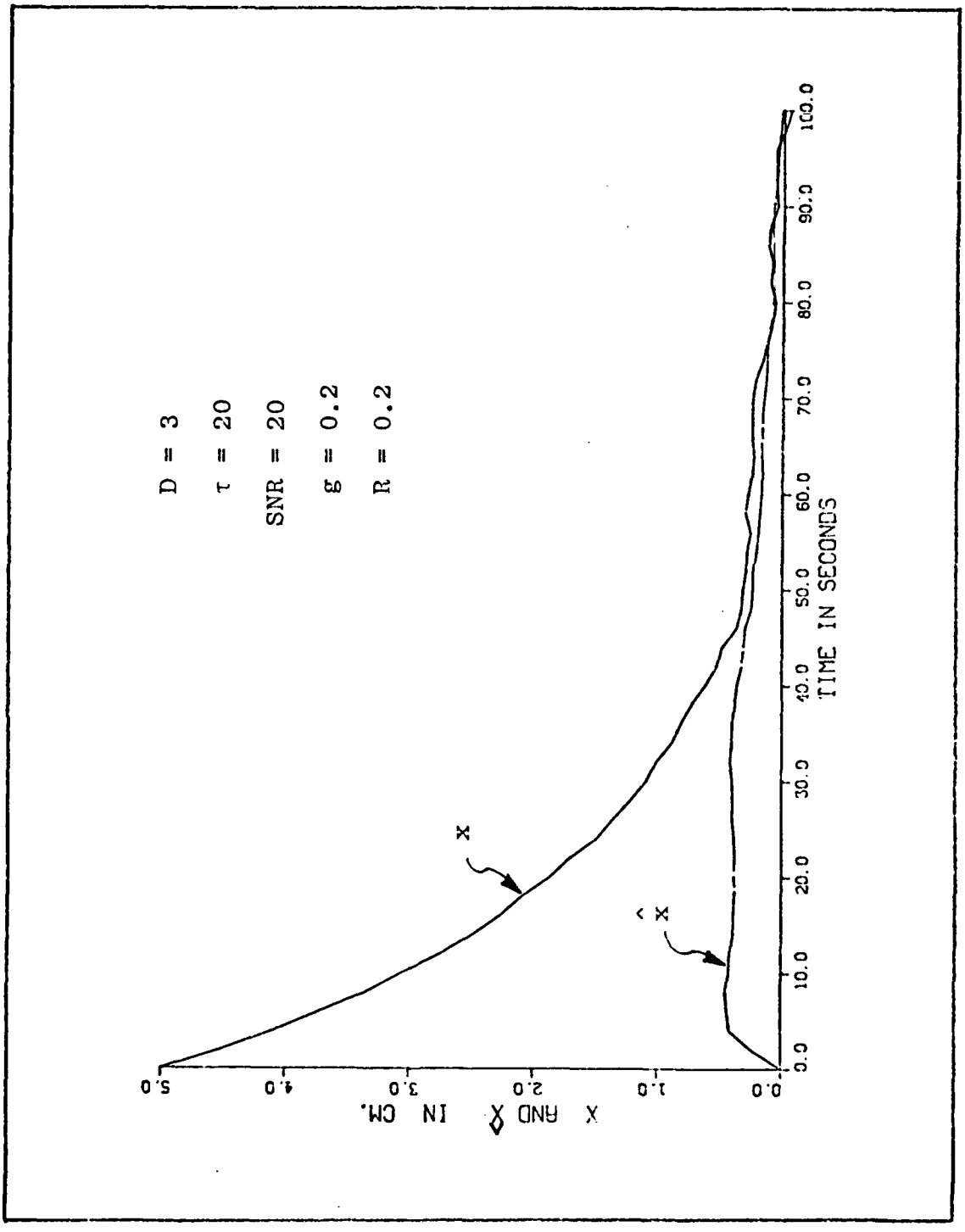


Figure 22. Unsuccessful Acquisition.

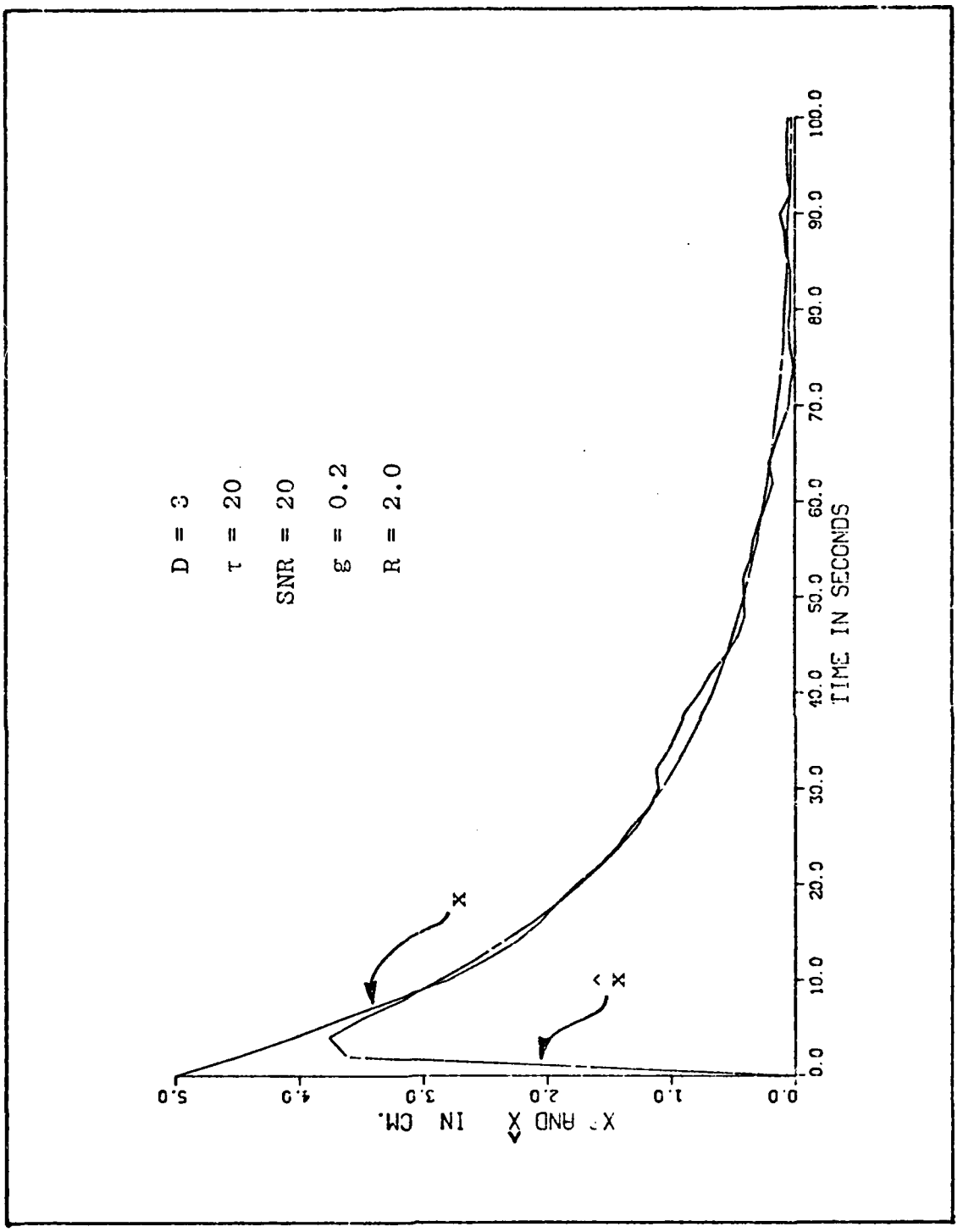


Figure 23. Successful Acquisition.

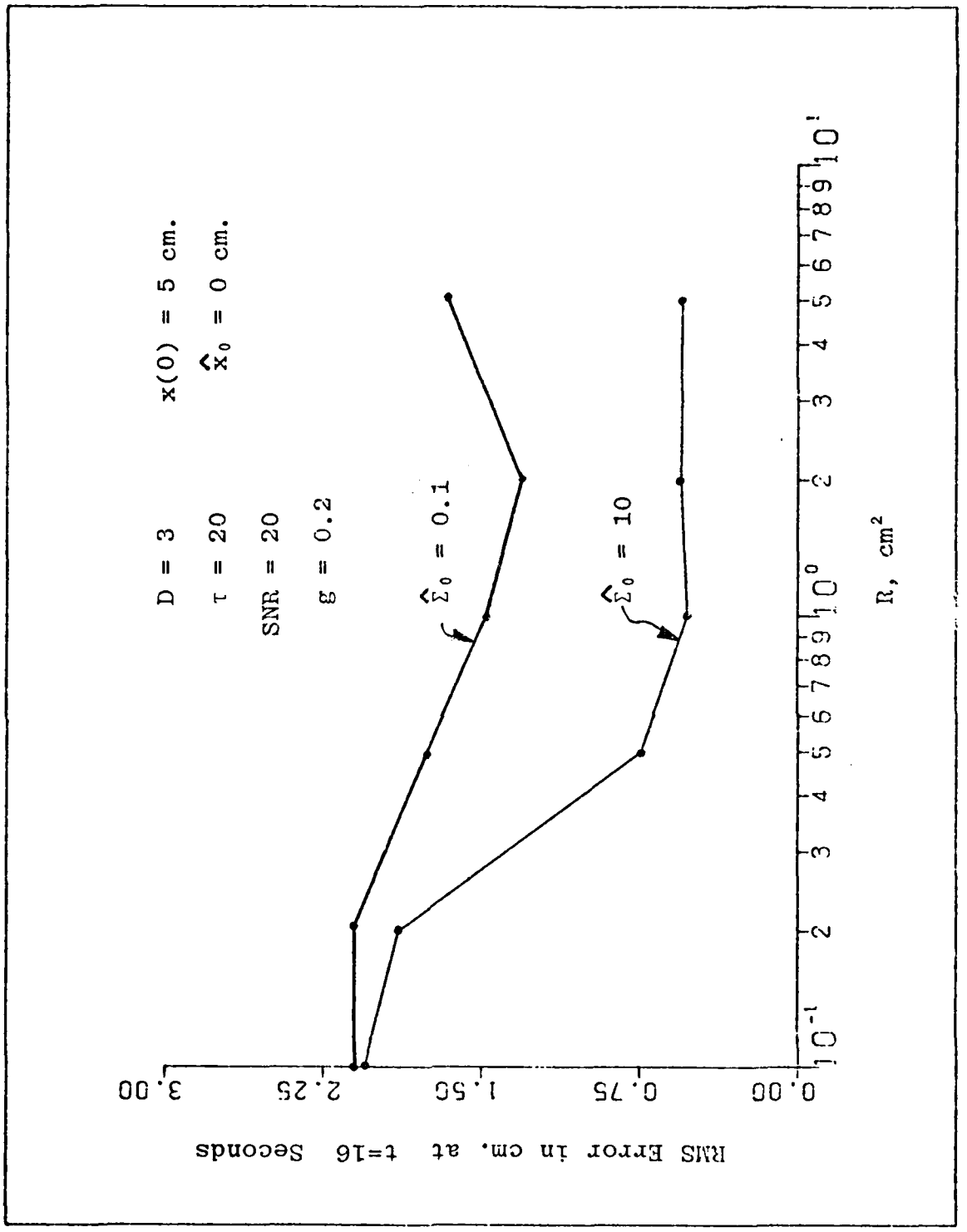


Figure 24. Estimator Acquisition Performance versus R.

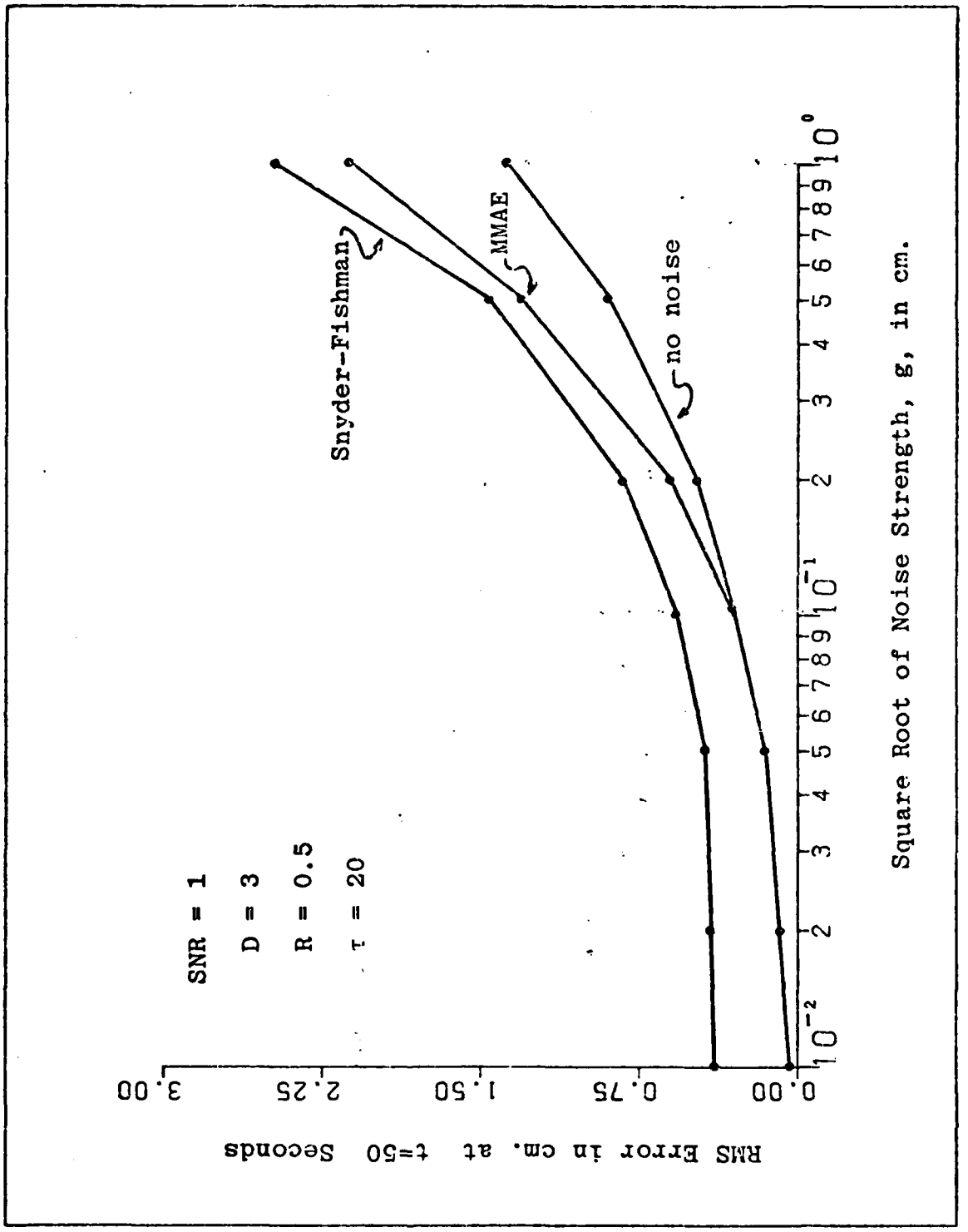


Figure 25. Performance Bounds

corresponding to the case in which the adaptive estimator perfectly deciphers which events are due to signal and which are due to noise. This was accomplished by setting the SNR = 10^6 in the input parameters to the simulation. Thus, there was almost never a noise event and the RMS error displayed in the lower curve is the best possible for the simulation parameter set shown.

An upper bound on performance was obtained by operating a single Snyder-Fishman filter against the same noisy input data for the same filter parameter set. Recall that the Snyder-Fishman filter accepts each observation as having been caused by the signal process.

As can be seen in Figure 25, the RMS error of the multiple model adaptive estimator (with the simplifications described in Chapter V) matches the lower bound on the error for values of $g \leq 0.1$ (square root of noise strength). Above this value, the RMS error of the simplified estimator is greater than the lower bound, but in all cases it is better than the performance of the Snyder-Fishman filter. A figure of merit which takes into consideration $\underline{F}(t)$, $\underline{G}(t)$, and $\underline{Q}(t)$ is the square root of the steady state covariance (SS/RMS) of $\bar{x}(t)$. For this one dimensional simulation example, the RMS value of the steady state variance is $\sqrt{\tau g^2/2}$. In terms of the SS/RMS, the RMS error of the multiple model adaptive estimator matches the lower bound for SS/RMS $\leq$ 0.3162, and the error is greater than the lower bound for larger values of SS/RMS. For the parameter

set shown in Figure 25, SS/RMS = 0.3162 for g = 0.1 and SS/RMS = 0.6324 for g = 0.2 . The significance of this is that as the SS/RMS value becomes larger than the "dispersion", R, of the signal rate parameter, the multiple model adaptive estimator's error is greater than the lower bound. For values of SS/RMS less than R , the estimator essentially makes perfect decisions as to which events were caused by signal and which were caused by noise.

Note that in the MMAE curve and the Snyder-Fishman curve, the expected signal to noise count ratio (SNR) is one; there is one expected noise event for each signal event.

VI.5 Other Simulation Considerations

As mentioned before, all of the simulation results shown in this chapter are based on knowledge of the true dynamical model of $\underset{\sim}{\bar{x}}(t)$ Although the performance levels will change, in general, with mismatched models, the trends displayed in these simulations should still be evident.

A model mismatch could easily arise for several reasons. For example, we may not know exactly how to model the dynamics of $\underset{\sim}{\bar{x}}(t)$ or we may wish to approximate a known model with a simpler model to reduce the computational load of the elemental filters.

One method to compensate for a mismatch is to add pseudonoise to the elemental filter models (Ref. 27 vol. 1:224). This is a common technique in Kalman filter tuning and has the effect of reducing the filter's confidence in

its own model. A second technique which may be useful is to impose an artificial lower bound on the elemental weighting factors (Refs. 2,27 vol.2). This would have the effect of putting more confidence in the less likely hypotheses; the overall estimate would be more heavily influenced by the elemental filter outputs associated with these less likely hypotheses. In general, this will tend to increase the error over that obtained when the model is known exactly; however, it may prevent a catastrophic failure in which the estimator "locks" onto a completely incorrect hypothesis model.

VI.6 Summary

In this chapter, a one dimensional model is specified for a tracking application in which there is no feedback control. Simulation results are presented based on the MMAE simplifications of Chapter V. The RMS error performance sensitivities to the major parameters of the estimator are presented and discussed.

The results are indicative of the overall RMS error performance of the multiple model adaptive estimator and of the performance trends as various parameters are varied. This simplified filter implementation shows excellent acquisition and tracking properties when the various filter parameters are tuned to the underlying process of interest. This is particularly notable considering the low data rate of the signal and the low expected signal to noise count ratios expected.

VII. Conclusions and Recommendations

VII.1 Conclusions

The goal of this research was to develop an estimator structure for the particle beam problem which is optimum, according to some appropriate criterion, and which is insensitive to point process noise corruption in the measurements. The multiple model adaptive estimator, presented in Chapter II, provides the minimum mean squared error estimate of the underlying process of interest. When the models are defined as separate sequence hypotheses, the estimator can reduce the effect of noise on the estimate. The development is valid for any point process signal in point process noise; it does not rely on the assumed conditionally Poisson statistics of the particle beam application which motivated this research. The full scale estimator does require an exponentially growing number of individual hypothesis filters.

The cross product space modeling concepts of Chapter III provide a means of calculating the individual filter weighting factors. This development is also valid for a general point process signal in point process noise application as long as the regularity conditions are met.

The specific analytical model for the particle beam application meets the regularity conditions. The cross product space modeling concepts allow feedback from the observations to the model, thus providing a way to define control inputs.

The specific expressions for the estimator suitable for the particle beam problem are developed in Chapter IV. The examples presented provide insight into the structure of the full scale estimator.

The simplifications to the full scale estimator presented in Chapter V result in a suboptimal filter which has greatly reduced requirements for calculation and storage. The approximation for the signal rate parameter estimate is appropriate for the particle beam problem when the model of the underlying process is known. The use of data windowing to stop the growth in the number of elemental hypothesis filters is applicable to the general point process multiple model adaptive estimator. Three methods are proposed to implement the data window concept. The "Best Half" method is recommended because it limits the growth, it implicitly includes information from the entire measurement history, and it is relatively simple to implement.

The simulation results of Chapter VI show that the suboptimal filter (using the simplifications developed in Chapter V) is extremely good at reducing the error caused by

point process noise. Performance trends are demonstrated as several of the parameters of the filter are varied. The performance degrades slowly, even at very low signal count rates and very low signal to noise count ratios. It is expected that the performance would degrade more rapidly if the estimator had inaccurate knowledge of the dynamics of the underlying process.

VII.2 Recommendations

There are several related areas of research which could provide immediately useful results. First, it is recommended that a method of predicting the error of the multiple model adaptive estimator be developed. Equation (193) provides the means of calculating the covariance of the full-scale multiple model adaptive estimator; however, the filter covariances of the elemental Snyder-Fishman filters depend on the space-time observations, as do the calculations of the weighting factors. Both of these are required for calculation of the overall covariance of the multiple model adaptive estimator. Currently, only simulation techniques provide a means of determining performance.

The second area of recommended research is in the convergence of the estimator. The key question to be determined here is whether the full scale estimator converges to the correct hypothesis sequence, and how rapid

the convergence is. The information theoretic ideas of Baram (Ref. 6), and Hawkes and Moore (Refs. 15,16) appear to provide promising avenues of research.

Another area which could provide immediately useful results is in the definition of the stochastic optimal controller for this point process problem. The cross product space modeling ideas allow the necessary feedback control; however, the specific optimum form of the control is not addressed in this dissertation. A natural follow-on topic is the investigation of a separation theorem for the optimum controller, or perhaps the effects of forced certainty equivalence (Ref. 27 vol. III:17).

The final research suggested is to explore the effects of imperfect knowledge of the dynamics of the underlying process (model mismatch). As described in Chapter VI, the estimator simulations were implemented with perfect knowledge of the dynamical model. A mismatch would certainly degrade the performance of the estimator. The extent of the degradation is of great importance in cases where a simplified model is desired due to the complexity of the true model, or where there is a lack of accurate knowledge about the true model.

Bibliography

1. Aoki, M. Optimization of Stochastic Systems. New York: Academic Press, 1967.

2. Athans, M. and C. B. Chang. "Adaptive Estimation and Parameter Identification Using Multiple Model Estimation Algorithm." Technical Note. M.I.T. 1976-28, 23 Jun 1976. AD-A028510.

3. Athans, M., D. Castanon, K. Dunn, C. Greene, W. Lee, N. Sandell, and A. Willsky. "The Stochastic Control of the F-8C Aircraft Using the Multiple Model Adaptive Control (MMAC) Method-I: Equilibrium Flight," IEEE Trans. AC: 768-780 (Oct 77).

4. Athans, M., R. H. Whiting, and M. Gruber. "A Suboptimal Estimation Algorithm with Probabilistic Editing for False Measurements with Applications to Target Tracking with Wake Phenomena," IEEE Trans. AC 22-3: 372-384 (Jun 1977).

5. Athans, M. and D. Willner. "A Practical Scheme for Adaptive Aircraft Flight Control Systems." NASA-TN-D-7647, p. 315-336. MIT, 1973. AD 760790.

6. Baram, Y. Information, Consistent Estimation and Dynamic System Identification. Dissertation. Massachusetts Institute of Technology, Cambridge, Mass. Nov 1976.

7. Chang, C. B., and M. Athans. "State Estimation for Discrete Systems With Switching Parameters," IEEE Trans. AES, AES-14, No. 3: 418-424 (May 1978).

8. Chang, C. B., and K. P. Dunn. "On GLR Detection and Estimation of Unexpected Inputs in Linear Discrete Systems," IEEE Trans. AC 24-3: 499-501 (Jun 1979).

9. Deshpande, J. G., T. N. Upadhyay, and D. G. Lainiotis. "Adaptive Control of Linear Stochastic Systems," Automatica, 9: 107-115 (1973).

10. Egardt, B. "Stochastic Convergence Analysis of Model Reference Adaptive Controller," IEEE Conference on Decision and Control (1980): 1128-1130, Albuquerque, NM, (1980).

11. Fishman, P. M. Statistical Inference for Space-Time Point Processes. Dissertation. Washington University, St. Louis, Missouri. 1974.

12. Fishman, P. M. and D. L. Snyder. "The Statistical Analysis of Space-Time Point Processes," IEEE Trans. IT 22-3: 257-274 (May 1976).

13. Gagliardi, R., and S. Karp. Optical Communications. New York: Wiley, 1976.

14. Greene ,C. S. and A. S. Willsky. "An Analysis of the Multiple Model Adaptive Control Algorithm," IEEE Conference on Decision and Control (1980): 1142-1145, Albuquerque, NM, (1980).

15. Hawkes, R. M. and J. B. Moore. "Performance Bounds for Adaptive Estimation," Proceedings of the IEEE, 64, No. 8: 1143-1150 (Aug 1976).

16. -----. "Performance of Bayesian Parameter Estimators for Linear Signal Models," IEEE Trans. AC: 523-527 (Aug 1976).

17. Kazakos, D. "The Bhattacharyya Distance and Detection Between Markov Chains," IEEE Trans. IT, 24-6: 747-754 (Nov 1978).

18. Kolmogorov, A. N., and S. V. Fomin. Introductory Real Analysis, edited and translated by R. A. Silverman. New York: Dover, 1975.

19. Lainiotis, D. G. "Optimal Adaptive Estimation: Structure and Parameter Adaptation," IEEE Trans. AC-16: 160-170 Apr 1971.

20. -----. "Joint Detection, Estimation and System Identification," Information and Control 19: 75-92 (1971).

21. -----. "Sequential Structure and Parameter-Adaptive Pattern Recognition - Part I: Supervised Learning," IEEE Trans. IT, 16-5: 548-556 (Sep 1970).

22. ------. "Supervised Learning Sequential Structure and Parameter Adaptive Pattern Recognition: Discrete Data Case," IEEE Trans. IT: 106-110 (Jan 1971).

23. Liporace, L. A. "Variance of Bayes Estimates," IEEE Trans. IT, 17-6: 665-669 (Nov 1971).

24. Ljung, L. "Analysis of Recursive Stochastic Algorithms," IEEE Trans. AC, 22-4: 551-575 (Aug 1977).

25. Macchi, O., and B. C. Picinbono. "Estimation and Detection of Weak Optical Signals," IEEE Trans. IT 18-5: 562-573 (Sep 1972).

26. Magill, D. T. "Optimal Adaptive Estimation of Sampled Stochastic Processes." IEEE Trans. AC 10-4: 434-439 (Oct 1965)

27. Maybeck, P. Stochastic Models, Estimation, and Control. New York: Academic Press, 1979 (vol. I), 1982 (vols. II and III).

28. McGarty, T. Stochastic Systems and State Estimation. New York: Wiley, 1974.

29. Moore, J. B. and R. M. Hawkes. "Decision Methods in Dynamic System Identification." IEEE Conference on Decision and Control (1975): 645-650, Houston, TX, (1975).

30. Morgan, A. P. "A Method of Estimating the Rate of Convergence of an Adaptive Observer," IEEE Conference on Decision and Control (1980): 1146-1151, Albuquerque, NM, (1980).

31. Murphy, D. J. "Batch Estimation of a Jump in the State of a Stochastic Linear System." IEEE Trans. AC: 275-276 (Apr 1977).

32. Neveu, J. Mathematical Foundations of the Calculus of Probability. San Francisco: Holden- y, Inc., 1965.

33. Papoulis, A. Probability, Random Variables, and Stochastic Processes. New York: McGraw-Hill, 1965.

34. Reiffen, B. and H. Sherman. "An Optimum Demodulator for Poisson Processes: Photon Source Detectors," Proceedings of the IEEE: 1316-1320 (Oct 1963).

35. Rhodes, I. B., and D. L. Snyder. "Estimation and Control Performance for Space-Time Point Processes," IEEE Trans. AC 22-3: 338-346 (Jun 1977).

36. Royden, N. L. Real Analysis. Second edition. New York: Macmillan Company, 1968.

37. Rohrs, C., L. Valavani, and M. Athans. "Convergence Studies of Adaptive Control Algorithms," IEEE Conference on Decision and Control (1980): 1138-1141, Albuquerque, NM, (1980).

38. Santiago, J. "Fundamental Limitations of Optical Trackers," Master's Thesis, AFIT, Wright-Patterson AFB, Ohio, 1978.

39. Sanyal P., and C. N. Shen. "Bayes' Decision Rule for Rapid Detection and Adaptive Estimation with Space Applications," IEEE Trans. AC: 228-231 (Jun 1974).

40. Sawaragi, Y., Sunahara, Y., and Nakamizo, T. Statistical Decision Theory in Adaptive Control System. New York: Academic Press, 1967.

41. Segall, A. "Optimal Control of Noisy Finite-State Markov Processes," IEEE Trans. AC 22-2: 179-186 (Apr 1977).

42. Segall, A., and T. Kailath. "The Modeling of Randomly Modulated Jump Processes," IEEE Trans. IT 21-2: 135-143 (Mar 1975).

43. Sims, F. L. and D. G. Lainiotis. "Recursive Algorithm for the Calculation of the Adaptive Kalman Filter Weighting Coefficients." IEEE Trans. AC: 215-218 (April 1969).

44. Smith P., and G. Buechler. "A Branching Algorithm for Discriminating and Tracking Multiple Objects," IEEE Trans. AC: 101-104 (Feb 1975).

45. Snyder, D. L. "Filtering and Detection for Doubly Stochastic Poisson Processes," IEEE Trans. IT 18-1: 91-102 (Jan 1972).

46. ------. Random Point Processes. New York: John Wiley and Sons, Inc., 1975.

47. Snyder, D. L. and P. M. Fishman. "How to Track a Swarm of Fireflies by Observing their Flashes," IEEE Trans. IT: 692-695 (Nov 1975).

48. Snyder, D. L., I. B. Rhodes, and E. V. Hoversten. "A Separation Theorem for Stochastic Control Problems with Point-Process Observations," Automatica 13: 83-87 (1977).

49. Vaca, M. V., and S. A. Tretter. "Optimal Estimation for Discrete Time Jump Processes," IEEE Trans. IT 24-3: 289-296 (May 1978).

50. Varaiya, P. "The Martingale Theory of Jump Processes," IEEE Trans. AC 20-1: 34-42 (Feb 1975).

51. Wenk, C. J. and Y. Bar-Shalom. "A Multiple Model Adaptive Control Algorithm for Stochastic Systems with Unknown Parameters." IEEE Conference on Decision and Control Proc.: 723-730, Fort Lauderdale, FL, (1979).

52. Willsky, A. S. "A Survey of Design Methods for Failure Detection in Dynamic Systems." Automatica, 12: 601-611 (1976).

53. Willsky, A. S., and H. L. Jones. "A Generalized Likelihood Ratio Approach to the Detection and Estimation of Jumps in Linear Systems," IEEE Trans. AC: 108-112 (Feb 1976).

54. Zacks, S. The Theory of Statistical Inference. New York: John Wiley and Sons, Inc., 1971.

Appendix A

In Chapter III, a cross product space model is developed and regularity results from Fishman (Ref. 11) are used to provide the means of calculating the weighting factors for the multiple model adaptive estimator. As described previously, this approach results in calculation of the elemental estimates through the use of the Snyder-Fishman filter and a simple weighted sum to calculate the overall estimate. The penalty of this method is the exponential growth of memory and calculation time requirements for the full-scale estimator.

The regularity results of Chapter III, along with an Itô diffusion differential rule, allow direct estimation of the process $\underset{\sim}{\bar{x}}(t)$. By "direct estimate," we mean an expression for $\underset{\sim}{\hat{\bar{x}}}(t)$ (perhaps in differential form) which does not use multiple model adaptive methods. The advantage of this is that there is no exponential growth in memory or calculation time requirements due to an expanding number of hypotheses, as in the multiple model estimator. The disadvantage to the direct method is that it requires evaluation of integrals which (depending on the specific system model) can be much more complicated than those

required by the multiple model method.

We begin the development by letting $\bar{V}$ be a regular space-time point process

$$\bar{V}:[t_0,t)X\Omega\rightarrow[t_0,t)XR^m$$

which satisfies Theorem III-1. Let $(\Omega,B,P(\cdot;\omega_s))$ be a probability space as defined in Chapter III and let (Ω_s,A_s,P_s) be a probability space defined as the cross product of two individual probability spaces

$$\Omega_s = \Omega_{s_1} X \Omega_{s_2}$$

where Ω_{s_2} is as defined in Chapter III (equation 99) and models the randomness of the noise process and Ω_{s_1} models the randomness of the signal process. Let A_{st} be the subsigma field of events generated for $t\varepsilon[t_0,t)$.

For the beam pointing and tracking problem, we are interested in estimating the process $\bar{x}(t;\omega;\omega_{s_1})$ given observations of the point process $\bar{V}$, that is, in generating

$$\hat{\bar{x}}(t) \triangleq E\{\bar{x}(t;\omega;\omega_{s_1})|\mathcal{B}_t\} \qquad \text{(A-1)}$$

where the signal rate parameter is dependent on $\underset{\sim}{\bar{x}}(t)$ as in equation 3.

We model the process of interest as the Itô diffusion process

$$d\bar{x}(t;\omega;\omega_{s_1})=\bar{a}(t;\omega;\omega_{s_1})dt+\underline{b}(t;\omega;\omega_{s_1})d\bar{u}(t;\omega_{s_1}) \qquad \text{(A-2)}$$

where $\bar{u}$ is a k dimensional Wiener process and $\bar{a}$,$\underline{b}$, and $\bar{u}$ are $\mathcal{A}_{st}\otimes\mathcal{B}_t$ measurable processes. This is a more general model than that given in equation 4, and we assume that a unique solution to equation A-2 exists.

For this Itô model, Fishman (Ref. 11) has shown the following differential rule. For a proof, see reference 11:170-174.

Theorem A-1, Generalized Ito Differential rule. Let

$$d\bar{\zeta}(t;\omega;\omega_s) = \bar{a}(t;\omega;\omega_s)dt+\underline{b}(t;\omega;\omega_s)d\bar{u}(t;\omega_s)$$

$$+\int_Y \bar{\gamma}(t,\bar{\pi};\omega;\omega_s)N(dt\times d\bar{\pi};\omega;\omega_s) \qquad \text{(A-3)}$$

$$Y \subset R^n$$

where $\bar{a}, \underline{b}$, and $\bar{u}$ are as in equation A-2, and $\bar{\gamma}$ is an m dimensional $A_{st}\otimes B_t$ – measurable process which is left continuous in t and continuous in $\bar{r}$. If $\psi(t,\bar{\xi}(t))$ has a continuous derivative in t and continuous partial derivatives of second order in the components $\xi_1, \xi_2, \ldots \xi_m$ of $\bar{\xi}$ for $t\varepsilon[t_0,T), \bar{\xi}\varepsilon R^m$ then (w.p.1)

$$d\psi(t,\bar{\zeta}(t)) = \frac{\partial\psi}{\partial t}\,dt + \left\langle \frac{\partial\psi}{\partial\bar{\zeta}}\,,\ \bar{a}(t)\right\rangle dt$$

$$+ \tfrac{1}{2}\ \text{trace}\left[\underline{b}\underline{b}^{T}\ \frac{\partial^{2}\psi}{\partial\bar{\zeta}^{2}}\right]dt + \left[\frac{\partial\psi}{\partial\bar{\zeta}}\right]^{T}\underline{b}\,d\bar{u}(t) \qquad \text{(A-4)}$$

$$+\int_{Y}\left[\psi(t,\bar{\zeta}(t)+\bar{\gamma}(t,\bar{\pi}))-\psi(t,\bar{\zeta}(t))\right]N(dt\,X\,d\bar{\pi})$$

The angle brackets in equation A-4 denote the vector inner product.

In order to use this differential rule to obtain the estimate $\hat{\bar{x}}(t)$ of the process $\bar{x}(t)$, let $\bar{x}(t)$ be defined as in equation A-2 and let

$$\eta(t;\omega;\omega_s) \triangleq -\int_{t_0}^{t}\int_{Y}[\phi(\tau,\bar{\pi};\omega;\omega_s)-\hat{\phi}_\tau(\tau,\bar{\pi};\omega)]d\bar{\pi}d\tau$$

$$+\int_{t_0}^{t}\int_{Y}\ln\left[\frac{\phi(\tau,\bar{\pi};\omega;\omega_s)}{\phi_\tau(\tau,\bar{\pi};\omega\ )}\right]N(d\tau X d\bar{\pi}) \qquad \text{(A-5)}$$

where $\phi(t,\bar{r};\omega;\omega_s)$ is the hazard function for the regular space-time point process $\bar{V}$ and $\hat{\phi}_\tau \triangleq E\{\phi|\mathcal{B}_\tau\}$. We define $\bar{\zeta}(t)$ in the differential rule as

$$\bar{\zeta}(t) \triangleq \begin{bmatrix} \bar{x}(t) \\ \hline \eta(t) \end{bmatrix} = \begin{bmatrix} x_1(t) \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \hline x_n(t) \\ \hline \eta(t) \end{bmatrix} \tag{A-6}$$

With this definition for $\bar{\zeta}(t)$ we can now write equation A-3 as

$$d\bar{\zeta}(t;\omega;\omega_s) = \bar{a}'(t;\omega;\omega_s)dt + \underline{b}'(t;\omega;\omega_s)d\bar{u}(t;\omega_s) \tag{A-7}$$

$$+\int_Y \bar{\gamma}'(t,\bar{\pi};\omega;\omega_s)N(dt\chi\, d\bar{\pi};\omega;\omega_s)$$

where

$$\bar{a}'(t;\omega;\omega_s) \triangleq \begin{bmatrix} \bar{a}(t;\omega;\omega_s) \\ \hline -\int_Y (\phi-\hat{\phi})\,d\bar{r} \end{bmatrix} \tag{A-8}$$

$$\underline{b}'(t;\omega;\omega_s) \triangleq \begin{bmatrix} \underline{b}(t;\omega;\omega_s) \\ \hline 0\ 0\ \ldots\ 0 \end{bmatrix} \tag{A-9}$$

$$\bar{\gamma}'(t,\bar{r};\omega;\omega_s) \triangleq \begin{bmatrix} \bar{0} \\ \hline \ln\left[\frac{\phi}{\hat{\phi}}\right] \end{bmatrix} \tag{A-10}$$

and where the arguments on the hazard functions have been dropped for ease of notation.

For $i = 1,2, \ldots ,n$ we define

$$\psi_i(t,\bar{\zeta}(t)) \triangleq x_i(t)e^{\eta(t)} \tag{A-11}$$

and apply the differential rule, equation A-4, to obtain

$$d\psi_i(t,\bar{\zeta}(t)) = a_i(t)e^{\eta(t)}dt - x_i(t)e^{\eta(t)}\left[\int_Y (\phi-\hat{\phi})d\bar{\pi}\right]dt$$

$$+ \sum_{j=1}^{k} b_{ij}(t)e^{\eta(t)}du_j(t) \tag{A-12}$$

$$+ x_i(t)e^{\eta(t)}\int_Y (\phi-\hat{\phi})\hat{\phi}^{-1}N(dt\times d\bar{\pi})$$

From Fishman (Ref. 11:160) it can be shown that

$$\hat{x}_i(t) = E\{x_i(t)|\mathcal{B}_t\} = E_s\{x_i(t)e^{\eta(t)}\}$$

(A-13)

$$= E_s\{\psi_i(t,\bar{\zeta}(t))\}$$

We can apply the Fubini Theorem to equation A-13 to obtain

$$d\hat{x}_i(t) = E_s\{d\psi_i(t,\bar{\zeta}(t))\} \tag{A-14}$$

therefore, we can take the expectation of both sides of equation A-12 to arrive at

$$d\hat{x}_i(t) = \hat{a}_i(t)dt - \left[\int_Y \left\{\widehat{x_i(t)\phi} - \hat{x}_i(t)\hat{\phi}\right\} d\bar{\pi}\right]dt$$

$$+ \int_Y \left\{\widehat{x_i(t)\phi} - \hat{x}_i(t)\hat{\phi}\right\}\hat{\phi}^{-1}N(dt\times d\bar{\pi}) \tag{A-15}$$

In vector form, the differential representation of $\hat{\bar{x}}(t)$ is

$$d\hat{\bar{x}}(t) = \hat{\bar{a}}(t)dt - \left[\int_Y \left\{\widehat{\bar{x}(t)\phi} - \hat{\bar{x}}(t)\hat{\phi}\right\} d\bar{\pi}\right] dt$$

$$+ \int_Y \left\{\widehat{\bar{x}(t)\phi} - \bar{x}(t)\hat{\phi}\right\} \hat{\phi}^{-1} N(dt \times d\bar{\pi}) \qquad \text{(A-16)}$$

$$\hat{\bar{x}}(t_0) = E\{\bar{x}(t_0)\}$$

The conditional error covariance matrix

$$\hat{\underline{\Sigma}}_t \triangleq E\left\{[\bar{x}(t) - \hat{\bar{x}}(t)][\bar{x}(t) - \hat{\bar{x}}(t)]^T \mid B_t\right\}$$

(A-17)

$$= \widehat{\bar{x}(t)\bar{x}^T(t)} - \hat{\bar{x}}(t)\hat{\bar{x}}^T(t)$$

can be expressed in differential form (Ref. 11:178) as

$$d\hat{\underline{\Sigma}}_t = \left[\left(\widehat{\bar{a}\ \bar{x}^T} - \hat{\bar{a}}\ \hat{\bar{x}}^T\right)+\left(\widehat{\bar{x}\ \bar{a}^T} - \hat{\bar{x}}\ \hat{\bar{a}}^T\right)+ \underline{b}\ \underline{b}^T\right]dt$$

$$+ \int_Y E\left\{(\bar{x}-\hat{\bar{x}})(\bar{x}-\hat{\bar{x}})^T(\phi-\hat{\phi})\,|\,\mathcal{B}_t\right\}\hat{\phi}^{-1}N(dt \times d\bar{\pi})$$

$$- \int_Y E\left\{(\bar{x}-\hat{\bar{x}})(\phi-\hat{\phi})\,|\,\mathcal{B}_t\right\}$$

$$\cdot\ E\left\{(\bar{x}-\hat{\bar{x}})^T(\phi-\hat{\phi})\,|\,\mathcal{B}_t\right\}\hat{\phi}^{-2}N(dt \times d\bar{\pi}) \qquad \text{(A-18)}$$

Equations A-16 and A-18 give the differential representation of the quantity to be estimated (directly) and the error covariance, respectively. We can now cast the particle beam model in terms of equation A-2 to obtain the direct estimator equations for this application.

As in Chapter III, the model for the process $\bar{x}(t)$ is given by

$$d\bar{x}(t;\omega;\omega_s) = \underline{F}(t)\bar{x}(t;\omega;\omega_s)dt + \underline{G}(t)d\bar{u}(t;\omega_s)$$

(A-19)

$$\underset{\sim}{\bar{x}}(t_0) = \underset{\sim}{\bar{x}}_0 \qquad t_0 \leq t$$

and the hazard function for the observed point process is

$$\phi(t,\bar{r};\omega;\omega_s) = \lambda_s(t,\bar{r};\omega;\omega_s) + \lambda_n(t,\bar{r};\omega;\omega_s) \qquad \text{(A-20)}$$

where the signal rate parameter is given by equation 103. The noise rate parameter is left general in form; we only require that the observed point process satisfy Theorem III-1 (it is regular).

By comparing our particle beam problem model (equation A-19) and the general Itô model (equation A-2), we can express the Itô model terms as

$$\bar{a}(t;\omega;\omega_s) = \underline{F}(t)\bar{x}(t;\omega;\omega_{s_1}) \tag{A-21}$$

$$\underline{b}(t;\omega;\omega_s) = \underline{G}(t) \tag{A-22}$$

We can substitute equations A-21 and A-22 into equation A-16 to obtain the differential expression for the direct estimate of $\hat{\bar{x}}(t)$ (for the particle beam model presented):

$$\hat{d\bar{x}}(t) = \underline{F}(t)\hat{\bar{x}}(t)dt$$

$$- \left[\int_Y \left\{ \widehat{\bar{x}(t)\phi} - \hat{\bar{x}}(t)\hat{\phi} \right\} d\bar{\lambda} \right] dt \tag{A-23}$$

$$+ \int_Y \left\{ \widehat{\bar{x}(t)\phi} - \hat{\bar{x}}(t)\hat{\phi} \right\} \hat{\phi}^{-1} N(dt \times d\bar{\lambda})$$

where

$$\hat{\bar{x}}(t) \triangleq E\left\{\bar{x}(t;\omega;\omega_{s_1}) \mid \mathcal{B}_t\right\} \tag{A-24}$$

$$\hat{\phi} \triangleq E\left\{\lambda_s(t,\bar{r};\omega;\omega_{s_1}) + \lambda_n(t,\bar{r};\omega;\omega_{s_2}) \mid \mathcal{B}_t\right\} \tag{A-25}$$

$$\widehat{\bar{x}(t)\phi} \triangleq E\left\{\bar{x}(t;\omega;\omega_{s_1})\phi(t,\bar{r};\omega;\omega_s) \mid \mathcal{B}_t\right\} \tag{A-26}$$

Conditional expected values of the form of equations (A-23) through (A-26) are difficult to solve. One approach is to assume a density for the quantity in question and then solve for the moments (or a partial set of the moments) (Ref. 27 Vol.II). The complexity of solving for either directly or through the use of an assumed density motivates the multiple model adaptive estimator approach taken in Chapter II.

VITA

David E. Meer was born on 15 April 1949 in Iowa City, Iowa. He attended the University of Iowa in Iowa City and graduated with a B.S.E.E. in June, 1972. Upon graduation, he received a commission in the USAF through AFROTC. Prior to active duty, he was an Electronic Engineer for the University of Iowa Physics Department on their Helios, IMP, and Hawkeye satellites. After entering active duty, his first assignment was to the Communications Electronics Engineering School at Keesler AFB, Mississippi. Upon completion, he served as Maintenance Supervisor for the AN/FSS-7 Sea Launched Ballistic Missile Detection and Warning Radar for Detatchment 4, 14th Missile Warning Squadron, Mt. Laguna AFS, California. In May 1975, he was assigned to the 1961st Communications Group, Clark AB, Republic of the Philippines. He served as Tactical Communications Engineer and head of the Mobile Communications Division until being assigned to the School of Engineering, Air Force Institute of Technology in June 1978. He received his M.S.E.E. from AFIT in December 1979 and entered the Ph.D. program at AFIT in January 1980.

Permanent address: 1305 Franklin
Iowa City, Iowa 52240